INVESTIGACION Y CIENCIA

Edición en español de **SCIENTIFIC AMERICAN**



ARROZ

Marzo 1984

350 PTAS.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Número 90
Marzo 1984

INVESTIGACION Y CIENCIA *Edición en español de* SCIENTIFIC AMERICAN

EDITA: Prensa Científica, S. A.
Calabria, 235-239
Barcelona-29 (ESPAÑA)

Colaboradores de este número:

Asesoramiento y traducción:

Fernando José Rodilla: *Respuesta precoz de contragolpe*; Manuel Puigcerver: *Efectos atmosféricos de El Chichón*; Ramón Pascual: *Colisiones a alta energía entre núcleos atómicos*; Enrique Sánchez Monge: *Arroz*; Joandomènec Ros: *Lagartijas unisexuales*; Santiago Torres: *Control de la síntesis de ribosomas*; Luis Bou: *Empaquetamiento de esferas* y *Juegos de ordenador*; J. Vilardell: *Taller y laboratorio.*

Ciencia y sociedad:

Antonio Prevosti y José A. Cabezas Fernández del Campo

Libros:

Domingo Fletcher, Luis Bru, Juan A. Navarro, Jaume Nubiola y M. García Velarde



LA PORTADA

La pintura de portada muestra un agricultor asiático plantando arroz en un campo cubierto por unos 10 centímetros de agua. Característica distintiva del arroz es que, gracias a su sistema de ventilación, que lleva oxígeno de las hojas a las raíces, puede desarrollarse en suelos inundados. Algunas especies crecen hasta 25 centímetros diarios en suelos anegados y sobreviven en aguas de hasta cinco metros de profundidad (véase "Arroz", de M. S. Swaminathan, en este número). El arroz pertenece al género *Oryza*; su especie asiática, la que presenta mayor número de variedades, es *O. sativa.* Cerca de 2000 millones de personas fundan en el arroz la fracción principal de su dieta. El terreno dedicado al cultivo del cereal alcanza los 145 millones de hectáreas, pese a lo cual, de no haber sido por el reciente desarrollo de variedades de alto rendimiento y con resistencia a enfermedades, la demanda habría superado la producción mundial.

Suscripciones:

Prensa Científica, S. A.
Calabria, 235-239
Barcelona-29 (España)
Teléfono 322 05 51 ext. 37

Condiciones de suscripción:

España:
Un año (12 números): 3.850 pesetas

Extranjero:
Un año (12 números): 52 U.S. $

Ejemplar atrasado ordinario:
350 pesetas

Ejemplar atrasado extraordinario:
450 pesetas

Distribución para España

Distribuciones de Enlace, S. A.
Ausias March, 49 - Barcelona-10

Distribución para los restantes países:

Editorial Labor, S. A.
Calabria, 235-239 - Barcelona-29

Publicidad:

Madrid:
Gustavo Martínez Ovin
Avda. de Moratalaz, 137 - Madrid-30
Tel. 430 84 81

Cataluña:
Lourdes Padrós
Manuel Girona, 61, 2.° - Barcelona-34
Teléfono 204 45 83

Controlado por O.J.D.

PROCEDENCIA DE LAS ILUSTRACIONES

Pintura de la portada de Teresa Fasolino

Página	Fuente	Página	Fuente
9-17	Jerome Kuhl	67-72	Patricia J. Wynne
23	René Canul, Comisión Federal de Electricidad (*arriba*); Brian R. Wolff (*abajo*)	75-82	Ian Worpole
24	Brian R. Wolff	83	Michael Beer, Universidad Johns Hopkins, Theo Koller, Instituto Federal Suizo de Tecnología, y Masayasu Nomura, Universidad de Wisconsin
25-26	Todd Pink	84-85	Ian Worpole
27	Michael Matson, National Oceanic and Atmospheric Administration	88-92	Gabor Kiss
28	Arlin J. Krueger, Goddard Space Flight Center	93-94	N. J. A. Sloane, Laboratorios Bell, y Gabor Kiss
29	Todd Pink	96	John Horton Conway, Universidad de Cambridge, N. J. A. Sloane, Laboratorios Bell, y Gabor Kiss
30	Alan Robock, Universidad de Maryland, Michael Matson, National Oceanic and Atmospheric Administration y Todd Pink	97	N. J. A. Sloane, Laboratorios Bell, y Gabor Kiss
32	Todd Pink	98	John Leech, Universidad de Stirling, y Gabor Kiss
35-36	Lawrence Berkeley Laboratory	101	A. Blanco Freijeiro
37-39	Walken Graphics	102-104	Juan Cotoner y A. Blanco
40	Lawrence Berkeley Laboratory	105	A. Blanco
41-42	Walken Graphics	106	Juan Cotoner y A. Blanco
52	Jurgen Schmitt, The Image Bank	107-109	A. Blanco
54	Patricia J. Wynne	117	Jearl Walker (*arriba y abajo*), David A. Katz (*en el centro*)
55-57	Andrew Tomko	118-122	Michael Goodman
58	Charles Moore, Black Star (*arriba*); Andrew Holbrooke, Black Star (*abajo*)		
59	Andrew Tomko		
60	Patricia J. Wynne		
61-62	Andrew Tomko		

ISSN 0210-136X
Dep. legal: B. 38.999-76
Fotocomposición Tecfa
Pedro IV, 160 - Barcelona-5
Fotocromos reproducidos por GINSA, S.A.
Imprime GRAFESA
Gráfica Elzeviriana, S. A.
Napoles, 249 - Tel. 207 40 11
Barcelona-13
Printed in Spain - Impreso en España

Los autores

JOHN STEINBRUNER ("Respuesta precoz de contragolpe") es director de estudios de política exterior de la Institución Brookings. Licenciado por la Universidad de Stanford, se doctoró en ciencias políticas por el Instituto de Tecnología de Massachusetts (MIT), donde por algún tiempo ejerció la enseñanza. Luego se trasladaría al Instituto de Política de la Universidad de Harvard, continuando con las clases al tiempo que dirigía el Programa para la Ciencia y los Asuntos Internacionales. Ingresó más tarde en la Universidad de Yale, de la que fue profesor de ciencias políticas. Abandonó Yale en 1978 para responsabilizarse de su cargo actual.

MICHAEL R. RAMPINO y STEPHEN SELF ("Efectos atmosféricos de El Chichón") son geólogos que colaboran en un estudio sobre las erupciones volcánicas y su efecto sobre el clima mundial. Cuenta Rampino: "Me licencié por el Hunter College y obtuve el doctorado en ciencias geológicas por la Universidad de Columbia. Desde 1978 hasta 1980 disfruté de una beca de investigación otorgada por la Academia Nacional de Ciencias, para trabajar en el Instituto Goddard de Estudios Espaciales de la Administración Nacional de Aeronáutica y del Espacio (NASA). Desde 1980 soy investigador adjunto del Instituto Goddard, en el departamento de geología de Columbia. En Goddard me incorporé al grupo que trabajaba en la elaboración de un modelo del clima global; en concreto, uno que trataba de predecir los efectos atmosféricos de las erupciones volcánicas explosivas". Self, que nació en Inglaterra, comenzó su formación en la Universidad de Leeds; prosiguió sus estudios de geología en el Imperial College of Science and Technology, que le concedió el doctorado en esa materia. Realizó trabajos de campo en Nueva Zelanda y formó parte del cuerpo docente de la Universidad del estado de Arizona antes de acceder a su actual cargo en el departamento de geología de la Universidad de Texas en Arlington.

WM. C. McHARRIS y JOHN O. RASMUSSEN ("Colisiones a alta energía entre núcleos atómicos") son químicos nucleares que comparten un mismo interés por el tema de su artículo. McHarris estudió en el Oberlin College y en la Universidad de California en Berkeley; se doctoró en química nuclear por Berkeley en 1965, incorporándose entonces a la Universidad estatal de Michigan, donde enseña química y física. Rasmussen se licenció por el Instituto de Tecnología de California recibiéndose luego de doctor por Berkeley, institución en cuyo claustro ingresó en 1952 y al que pertenece desde entonces, salvo un intervalo: el cuatrienio (1969-1973) que le contrató la Universidad de Yale. Ha sido profesor visitante del Instituto Nobel de Física de Estocolmo y becario de la Fundación Nacional para la Ciencia del Instituto Niels Bohr.

M. S. SWAMINATHAN ("Arroz") dirige el Instituto Internacional de Investigaciones Arroceras de Filipinas. En su carrera profesional ha combinado los estudios genéticos con la aplicación de los resultados científicos a la práctica agrícola. Nacido en Kumbakonum, ciudad del sur de la India, se licenció por la Escuela Agrícola de Coimbatore, prosiguiendo estudios de genética en el Instituto de Investigaciones Agrarias de la India (IIAI). Tras una estancia en Holanda y en la Universidad de Cambridge, se recibió de doctor por esta institución en 1952. Ingresó entonces en la Universidad de Wisconsin. En 1954 regresó a la India, trabajando para el IIAI. En 1972 fue nombrado director general del Consejo indio de Investigaciones Agrarias, cargo en el que se mantuvo hasta 1979, cuando accedió a la secretaría de desarrollo agrícola y rural del gobierno indio.

CHARLES J. COLE ("Lagartijas unisexuales") es conservador del departamento de herpetología del Museo Americano de Historia Natural de Nueva York. Se licenció por la Universidad Wesleyana, completando el segundo ciclo universitario en la de Kansas; es doctor por la Universidad de Arizona. En la investigación sobre los saurios unisexuales que se describe en este artículo contó con la ayuda de su hijo, Jeffrey A. El estudio ha recibido apoyo financiero de la National Geographic Society y de la National Science Foundation norteamericanas.

MASAYASU NOMURA ("Control de la síntesis de ribosomas") escribe: "Recibí mi educación durante los tristes años que siguieron a la segunda guerra mundial en Japón. Me especialicé en microbiología y química de la fermentación en un departamento de ciencia aplicada de la Universidad de Tokio. Elegí ese departamento porque pretendía encontrar trabajo en la industria, no porque pensara convertirme en científico. No obstante, después de seguir algunos cursos de ciencia aplicada, creció mi interés por la ciencia básica. Por aquella época empezaron a llegar los ecos del desarrollo de la biología molecular, campo que me entusiasmó, y en 1957 me trasladé a los Estados Unidos con una beca postdoctoral. Durante los tres años siguientes trabajé en los laboratorios de Sol Spiegelman, en la Universidad de Illinois, James D. Watson, en Harvard, y Seymour Benzer, en Purdue. Volví al Japón en 1960 como profesor del Instituto de Investigación en Proteínas, de la Universidad de Osaka. A los tres años tomé la difícil decisión personal de abandonar mi país para ir a hacer ciencia a los Estados Unidos".

N. J. A. SLOANE ("Empaquetamiento de esferas") es miembro de la plantilla técnica de los Laboratorios Bell. Nacido en Gales, se formó en Australia y en los Estados Unidos. Se licenció por la Universidad de Melbourne, doctorándose por Cornell. Nombrado profesor auxiliar de ingeniería eléctrica de la Universidad de Cornell, en 1967, permaneció en ella por dos años, trasladándose después a los Laboratorios Bell. En su trabajo científico, Sloane se ha ocupado de teoría de codificación, teoría de comunicación y matemática combinatoria.

ANTONIO BLANCO FREIJEIRO ("Mineros y metalúrgicos antiguos en Huelva") es catedrático de arqueología de la Universidad Complutense de Madrid. Estudió filología y arqueología en las universidades de Madrid, Oxford y Heidelberg. Ha sido conservador del Museo del Prado y autor del *Catálogo de la escultura* de dicho museo; director de la Academia Española de Bellas Artes en Roma y catedrático de arqueología, epigrafía y numismática de la Universidad de Sevilla entre 1959 y 1973. Su libro de *Arte griego* va por la quinta edición y su *Arte antiguo del Asia Anterior* por la tercera; es autor, además, de varias obras sobre mosaicos romanos de España y colaborador como autor y traductor de muchos artículos aparecidos en *Investigación y Ciencia*. Es miembro del consejo asesor del IAMS (Institute of Archaeo-Metallurgical Studies, London), numerario de la Real Academia de la Historia en Madrid, del Deutsches Archäologisches Institut en Berlín y de otras sociedades eruditas.

Hace...

José M.ª López Piñero

... cuatrocientos años

Bartolomé Hidalgo de Agüero publicó *Avisos particulares de cirugía contra la común opinión*, donde expuso por vez primera el método de tratamiento de las heridas llamado "vía particular desecante" y otras críticas de las prácticas y de los fundamentos doctrinales de la cirugía tradicional.

La vida de Hidalgo de Agüero transcurrió por entero en la ciudad de Sevilla, desde su nacimiento en 1530 hasta su muerte el año 1597. En la capital andaluza hizo sus estudios y ejerció su profesión, en buena parte como cirujano del Hospital del Cardenal. La primera edición de los *Avisos* es un pliego impreso a doble columna, del que no se conocían ejemplares hasta que, en 1978, Esteban Torre localizó uno en el Archivo Municipal de Sevilla. Su texto fue incluido, junto con las demás obras de Hidalgo, en el volumen *Thesoro de la Verdadera Cirugía y Vía Particular contra la común*, publicado por su yerno en Sevilla en 1604, siete años después de su muerte, y reeditado más tarde en Barcelona (1624) y Valencia (1654).

La doctrina tradicional del "pus loable" –procedente del galenismo bajomedieval arabizado– justificaba la aplicación de curas emolientes de las heridas que, al "ablandar y madurar", favorecían la formación de pus con la intención de evacuar los supuestos humores alterados. La reacción contra esta práctica y su fundamento hasta el siglo XVI no fueron sometidos a una crítica en profundidad, una de cuyas principales manifestaciones fue la "vía particular" de Hidalgo.

La "vía particular" consistía fundamentalmente en la cura por "primera intención desecante" de las heridas en general, tanto incisas como contusas. Se oponía a la "vía común (que) cura humedeciendo con medicinas y con uso de manos e instrumentos de hierro". Hidalgo subrayó repetidamente su novedad, afirmando que "los antiguos no la hallaron", aunque intentó encontrar sus antecedentes en algunos textos de Galeno.

Especial interés tiene la forma en la que el cirujano sevillano demostró las ventajas de su método frente a la cura tradicional. Adujo para ello su experiencia durante dos décadas en el Hospital del Cardenal, empleando historias clínicas cuidadosamente recogidas, muchas de las cuales reprodujo en sus obras. Llegó incluso a comparar los resultados de ambos procedimientos a base de datos numéricos, lo que constituye uno de los más tempranos intentos de utilización de las estadísticas hospitalarias: "Cierta es esta vía particular –afirma– que llamamos primera intención desecante, respecto de la común vía y segunda intención humectante... y esto no lo puede entender, sino es quien haya sido curioso en la una y otra vía, habiendo curado y no poco tiempo por ambas y haya visto los provechos y daños de entrambas vías seca y húmeda. Y puedo afirmar como testigo de vista que he curado más de veinte años, y tenido lugares públicos, como es el Hospital del Cardenal de Sevilla, y otros, y mucho estudio, y concurso de estudiantes; y vistos los incómodos y daños y ejercitado la particular por la misma orden que la común, y he hallado tan grandes provechos que me han obligado, no sólo a ejercitarla, más también a comunicarla... Hice regular por el libro del dicho Hospital del Cardenal, donde asientan las heridas, y se ha visto que el año pasado de mil quinientos y ochenta y tres años, entraron cuatrocientos y cincuenta y seis, y murieron veinte. Y en dos meses y medio que hice asentar de por sí los heridos de cabeza, entraron cincuenta y siete, y salieron sanos cincuenta y murieron siete. Y en los años que han curado mis antecesores y yo por la común, se halla mayor número de los muertos que de los vivos".

Sobre la base de la "vía particular", Hidalgo se opuso a que se trepanase o legrase en la inmensa mayoría de las heridas cefálicas. Tanto su método curativo como esta actitud contraria a las intervenciones craneales fue mantenida por sus discípulos y seguidores, entre los que se encuentran algunos de los mejores cirujanos españoles de las primeras décadas del siglo XVII. Sobresale especialmente Pedro López de León y, con menor altura, Enrique Vaca de Alfaro y Pedro Gago Vadillo.

Los escritos de Hidalgo reunidos en el *Thesoro* se ocupan, además, de las cuestiones habituales en la literatura quirúrgica de la época. Aparecen los obligados estudios sobre heridas, úlceras y "apostemas" o tumoraciones, así como sobre fracturas y dislocaciones. En este último –una de las mejores exposiciones sobre el tema publicadas en la España de la época– el cirujano sevillano se muestra, en general, contrario al uso de aparatos para la reducción de luxaciones, e incluso de férulas para el tratamiento de las fracturas. Defiende una actitud moderada en el empleo de purgas y sangrías, e incluye un repertorio de medicamentos o "Antidotario" y capítulos o breves tratados acerca de la peste, el "tabardillo", la "modorra" y la sífilis.

Mención aparte merecen los textos anatómicos, en especial la "Anatome del cuerpo humano" y el estudio monográfico "De la historia del ojo". Ambos corresponden a un saber morfológico basado en la disección –que fue practicada asiduamente por Hidalgo– y son de orientación posvesaliana. Su posición contraria al empleo en cirugía de instrumentos que juzgaba inútiles la extiende también a la práctica de la disección. Uno de los puntos de los *Avisos* afirma que "para el uso de la anatome, solos tres instrumentos son menester, que son cuchillo y anzuelos y sierra; los demás son confusión, si son ferrales". Al desarrollar dicho punto en el *Thesoro* dice, entre otras cosas: "Los anatomistas, que han querido hacer el cortar dificultoso, començaron a cortar con tanta diferencia y confusión de instrumentos, los cuales son superfluos. Y que sea lo que digo verdad, yo he cortado y corto muchas veces, y no tengo necesidad de más de los tres instrumentos dichos... todos los que saben cortar saben esta verdad, si la quieren manifestar, como nosotros pretendemos en esto y en lo demás hacer notorio al mundo lo que tiene dificultad y lo que es fácil, para que los ánimos no estén suspensos con tanta confusión".

Hidalgo dedica capítulos duramente críticos a los errores del saber anatómi-

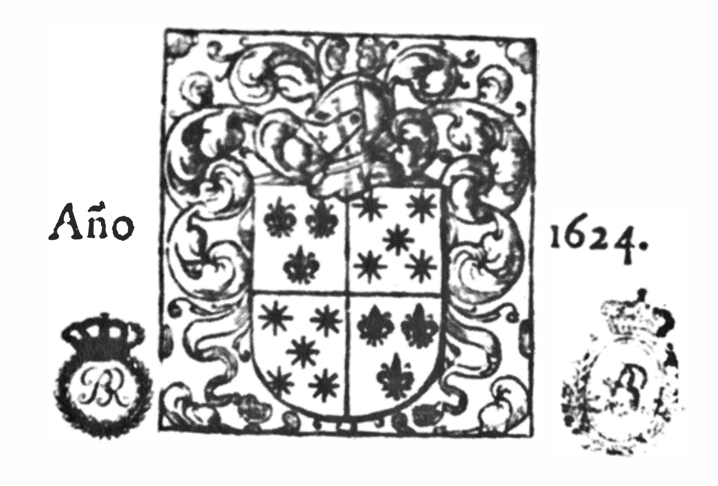
co tradicional, entre ellos, la presencia de *rete mirabile* en el cuerpo humano y la afirmación "que los nervios ópticos tienen cavidad manifiesta". Al ocuparse de la primera, dice que "los que han cortado cabezas de carneros o de vacas han hallado *rete mirabile*, mas los que han cortado estos sujetos y también hombres han visto la diferencia de lo uno y de lo otro y, por ser curiosos, reprehenden a los que, por haber hallado en carneros y en los demás brutos aquel plexo o *rete mirabile*, afirmaron que era también común a los hombres, el cual no se halla cortando en sujeto humano". En este contexto, elogia las indagaciones de cultivadores de la nueva anatomía, como el castellano Juan Valverde y el valenciano Luis Collado, al tiempo que "Galeno es reprehendido".

También critica las concepciones fisiológicas tradicionales incompatibles con los datos anatómicos. Por ejemplo, un punto de los *Avisos* dice: "Los que son negligentes en cortar cuerpos humanos afirman que el hueso sacro se abre al tiempo del parto y que después se torna a cerrar; lo cual es falso, porque ni antes ni después se abre". Más tarde, termina la demostración anatómica de este error, calificando a los que lo defendían de "hombres idiotas", como había hecho Vesalio.

THESORO
DE LA VERDA-
DERA CIRVGIA Y
VIA PARTICVLAR
contra la Comun.

COMPVESTO POR EL DOTOR Bartolome Hidalgo de Aguero Medico y Cirujano, con la qual se haze vn perfeto Cirujano.

Dirigido a Diego Hexarch Dotor en Medicina.



CON LICENCIA,

En Barcelona, Por Sebaſtian de Cormellas. *Y a ſu coſta.*

1. Portada de la edición barcelonesa del **Thesoro de la Verdadera Cirugía**, *volumen que recogió las obras de Bartolomé Hidalgo de Agüero, uno de los principales renovadores de la doctrina y la práctica quirúrgicas durante el siglo XVI.*



AVISOS PAR-
TICVLARES DE
CIRVGIA CONTRA LA CO-
mun opinion, Por el Doctor Bartolo-
me Hidalgo de Aguero Medi-
co, y Cirujano.

1 QVe ninguna herida de cabeça ſe ha de legrar, ni leuantar hueſſos, aunque punjan, y preman, y, eſtè intromiſſo el Craneo, ni ciſſo, ni fracto, ni coliſo.

2 Que todas las heridas de inſtrumento que corta, ſe han de aglutinar, y curar con deſecantes, y ni mas ni menos las contuſas piden la miſma curacion, quitandoles las coſas preternaturales.

3 Todas las heridas de cabeça, que fueren de inſtrumento que colide, y contunde, ſino touieren leſion en el Craneo, o fluxo de ſangre, no ſe han de formar.

4 Las heridas de la cabeça, que fueren de inſtrumento que colide, y tienen fractura, y hueſſos intromiſſos y pungentes, no ſe formaran deſpues del primer aparato con fuerte formacion ſino liuiana, y cada dia mas ligera, pues que no ſe han de legrar, ni trepanar, ni leuantar hueſſo, ni quitar los que pungen, y premen, ni hazer ningun vſo de manos de los que mandan los Autores.

¶¶ 4 5 Si

2. Página inicial de los **Avisos particulares de cirugía contra la común opinión**, *de Hidalgo de Agüero, relación de cincuenta puntos en los que criticaba prácticas y fundamentos teóricos de la cirugía tradicional.*

Respuesta precoz de contragolpe

Las presiones en favor de la adopción de esta estrategia, que contempla el disparo de armas nucleares al advertir el lanzamiento de las del enemigo, ponen de manifiesto la creciente inestabilidad del equilibrio nuclear

John Steinbruner

La decisión de la Administración Reagan de instalar 100 misiles MX en silos de Minuteman plantea serios interrogantes sobre su política en el campo de las fuerzas estratégicas nucleares. La teórica vulnerabilidad de la fuerza Minuteman a un ataque por sorpresa por parte de los misiles balísticos intercontinentales (ICBM) soviéticos, cada vez más precisos, fue el argumento esgrimido para justificar la prosecución del programa MX. Sin embargo, tras un detenido análisis de más de treinta modalidades alternativas de instalación de los MX, la Administración, siguiendo los consejos de la Comisión Presidencial para Fuerzas Estratégicas (la comisión Scowcroft), reconoció que ninguno de los modelos propuestos garantizaba la supervivencia de la fuerza MX ante un ataque nuclear soviético; ni podría, por tanto, reaccionar con una respuesta adecuada. Con semejante conclusión a la vista, lo menos que puede decirse de un plan que prevé el despliegue del primer contingente de los nuevos misiles MX en las viejas instalaciones es que le falta lógica.

La verdad es que la situación puede resultar aún peor. El MX de diez ojivas es un arma mucho más peligrosa que el Minuteman III, de menor tamaño y dotado sólo de tres ojivas, el más contundente de cuantos ICBM posee el arsenal norteamericano. Según las mejoras previstas en cuanto a precisión, la capacidad destructiva total de cada MX superará entre 16 y 24 veces la de un misil Minuteman III. Si en los demás aspectos no se produjese ninguna variación, el despliegue de los MX en silos de Minuteman incrementaría los alicientes de la Unión Soviética para intentar un ataque por sorpresa contra esas instalaciones, lo que a la postre menguaría la estabilidad del actual equilibrio de disuasión nuclear.

¿Se mantendrán sin alteración los demás aspectos? En su defensa de la última decisión sobre los MX, en abril de 1983, ante una Comisión del Congreso, el Secretario de Defensa, Caspar W. Weinberger, y el Presidente de la Junta de Jefes de Estado Mayor, General John W. Vessey, Jr., señalaron que algo podría cambiar, a saber, la disposición de Estados Unidos a soportar un ataque inicial antes de comprometerse a emprender una acción de represalia. Aunque ambas personalidades se negaron a entrar públicamente en detalles, de sus puntualizaciones se desprendía que Estados Unidos podría estar planeando acompañar el despliegue de misiles MX con una mayor inclinación al lanzamiento de sus ICBM en cuanto tuviese pruebas suficientes del inicio de un ataque soviético, antes de acusar el golpe.

Esta política, comúnmente denominada de "lanzamiento ante alerta" o de "respuesta precoz de contragolpe", la han defendido muchos estrategas como solución al problema de la vulnerabilidad de los ICBM; otros la han criticado por entenderla una peligrosa consecuencia del problema de la vulnerabilidad. Tras esa disparidad de criterios se esconde uno de los más difíciles problemas de la actual política de seguridad internacional.

Al abordar la cuestión, que guarda una relación directa con las posibilidades de una guerra nuclear, conviene distinguir varios niveles de análisis. En un primer nivel se esgrimen argumentos teóricos sobre estrategia nuclear, apoyados en cálculos estrictamente técnicos. En un segundo nivel se esbozan reflexiones prácticas, fruto de aplicar esas teorías a la estructura de la fuerza estratégica de Estados Unidos. Por último, existe la realidad militar subyacente, que raras veces se reconoce en los debates de la política de seguridad nacional. El problema que aquí se analiza cambia significativamente de carácter según se afronte desde cada una de esas diferentes perspectivas.

En el plano de la discusión teórica, la vulnerabilidad de los ICBM está clara y tajantemente definida. Por cálculos convencionales, los estrategas occidentales han evaluado la probabilidad de que una instalación fija de ICBM –un silo– sobreviva a un ataque nuclear. Estos cálculos se basan en un escaso número de factores técnicos, todos los cuales pueden medirse con cierto grado de precisión: la solidez de los silos, o resistencia al daño causado por una explosión, número, potencia explosiva, precisión y funcionamiento operativo conjunto de las ojivas atacantes. Las características de las ojivas más avanzadas hacen que toda proporción de ojivas contra silos superior a 2,5 a 1 dé una tasa de supervivencia de ICBM muy baja, próxima a la destrucción completa. Se estima que, en el momento actual, la Unión Soviética sobrepasa con creces esa proporción contra la fuerza ICBM estadounidense. Combinando 550 misiles Minuteman III, portadores de tres ojivas cada uno, con 100 misiles MX, portadores a su vez de diez ojivas, los Estados Unidos tendrían una capacidad semejante contra los 638 silos que albergan las más modernas ojivas de la fuerza ICBM soviética.

Según la teoría dominante de disuasión nuclear, que presupone que las acciones de ambas partes se rigen por refinados juicios racionales, esta vulnerabilidad mutua de las fuerzas ICBM crea una amenaza de fuerza nuclear limitada y la posibilidad de una coacción política fundada en tal amenaza. Es de suponer que, bajo la presión de una crisis grave, cualquier parte podría iniciar un ataque dirigido exclusivamente contra las instalaciones vulnerables de ICBM de la otra, al objeto de desarmarla parcialmente y establecer una superioridad militar. Al haber perdido lo más preciso y mejor dispuesto de su fuerza defensiva, la víctima, teóricamente racional, no podría responder de modo semejante y, ante las perspectivas de un daño aún mayor, renunciaría a emprender acciones de represalia

contra otros objetivos civiles y militares, buscando en consecuencia un entendimiento político que evitara la escalada bélica. Siguiendo esta lógica, la vulnerabilidad de ambos ICBM parece dar una victoria importante, aunque parcial, al bando que se apreste a explotar la situación.

La política sistemática de lanzamiento ante ataque (la aquí llamada respuesta precoz de contragolpe) solucionaría así ese problema. Cálculos técnicos, comparables en sus grandes rasgos científicos a los empleados para definir el problema de la vulnerabilidad de los ICBM, sugieren que, en principio, es detectable el inicio de un ataque del adversario y, factible, lanzar parte de la fuerza ICBM amenazada antes de recibir el golpe. La variante más refinada de esta idea, postulada por Richard L. Garwin, aboga por un sistema de sensores infrarrojos en el espacio para detectar los misiles atacantes en su fase de lanzamiento y medir la escala del ataque. La información se transmitiría a estaciones terrestres de los Estados Unidos y la confirmarían las correspondientes autoridades nacionales. Un sistema de comunicaciones de emergencia basado en un satélite ya situado en órbita o en un cohete especialmente lanzado a tal efecto retransmitiría entonces, de acuerdo con los planes establecidos, las órdenes de lanzamiento de una parte importante, aunque limitada, de la fuerza ICBM en peligro.

Según Garwin, habría un despliegue suficiente de sensores, enlaces de comunicación y procesamiento de la información del sistema para protegerlo contra un ataque directo y para hacer descender hasta niveles sumamente bajos la probabilidad teórica de una falsa detección de ataque y su respuesta violenta. En la mayoría de los casos, los mecanismos para el lanzamiento definitivo estarían bloqueados, dando así al sistema la posibilidad de una comprobación permanente del diagnóstico realizado, al tiempo que impediría –siempre en teoría– cualquier respuesta de represalia por un fallo debido al azar. Para casos imprevistos y urgentes por razones de seguridad, se instrumentaría como alternativa una serie de controles internos, que permitiría el funcionamiento del sistema de modo automático –tras la previa confirmación del ataque por los responsables de la decisión– o con diversos grados de intervención humana, para así proporcionar una protección más y evitar una acción de represalia no deseada.

En el marco teórico descrito, semejante sistema podría presentar al potencial agresor la suficiente incertidumbre para eliminar todo el incentivo de iniciar un ataque que ofrece la vulnerabilidad de las instalaciones de ICBM enemigas. Puesto que cualquier probabilidad razonablemente elevada de éxito en una respuesta precoz ejercería seguramente ese efecto, la víctima en potencia podría aceptar, digamos, un 20 por ciento de posibilidades de fracaso al emprender una acción de represalia contra ataques reales y mantener aún un fuerte poder disuasorio. De acuerdo con las teorías dominantes avaladas por los cálculos técnicos, si aceptamos que los ICBM son vulnerables, no parece existir razón alguna para descartar la solución que supone un sistema elaborado de respuesta precoz de contragolpe.

En un plano de análisis más práctico, existen buenas razones para dudar de la validez tanto del planteamiento del problema como de la solución al mismo. Los métodos organizativos por los que se rige el manejo de las armas estratégicas hacen muy improbable que un ataque por sorpresa contra silos de ICBM, así como la respuesta protectora de lanzamiento ante ataque, se desarrollen de acuerdo con los cálculos teóricos. Pero en la política estratégica de los Estados Unidos han pesado otras razones. Mientras que la cuestión de la vulnerabilidad de los silos se consideraba un auténtico problema que exigía solución práctica, la instrumentación segura de medidas de respuesta se juzgaba tan poco práctica que ni siquiera se intentó un esfuerzo sistemático para desarrollar la capacidad necesaria.

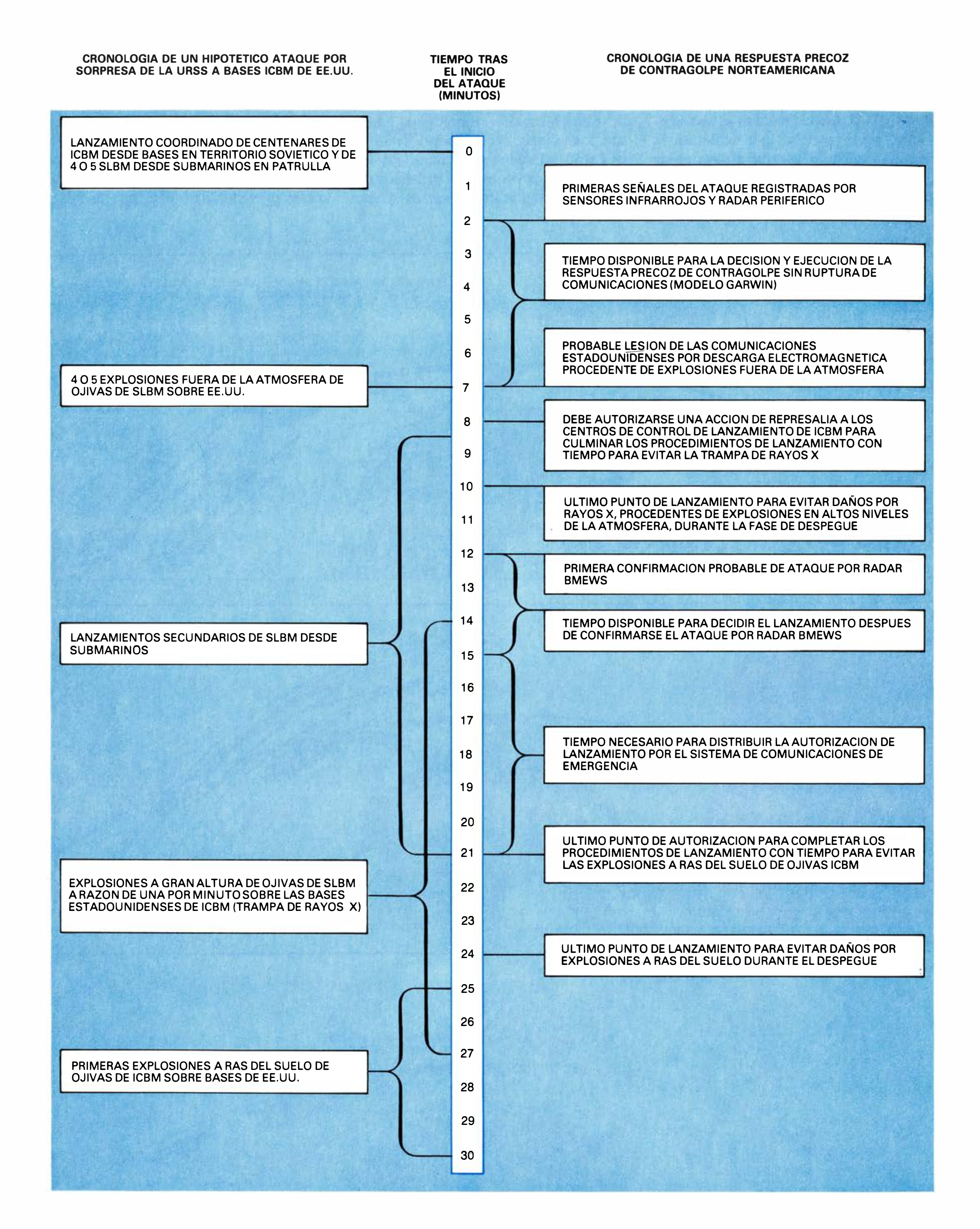
Los problemas prácticos del ataque y de la respuesta arrancan de un mismo origen: el conflicto entre las exigencias de paz y las de guerra, que puede contemplarse en términos técnicos como el contrapeso de dos tipos de errores. Al desplegar fuerzas ofensivas que disuadan de una guerra, el objetivo primordial en tiempos de paz es necesariamente evitar el uso accidental o no autorizado de tales armas (lo que los estados mayores denominan errores de tipo I). Evitar errores de ese tipo significa mantener un control negativo. Si llegase realmente a estallar la guerra, los ataques de represalia para lograr algún éxito militar importante se convertirían en el objetivo primordial; cobraría especial relieve entonces reducir los fallos en el lanzamiento de armas contra objetivos previamente asignados (errores de tipo II). La prevención de errores de este tipo exige un control positivo (activo). Por su propia naturaleza, los controles positivo y negativo se hallan en mutuo conflicto. Privilegiar uno supone postergar el otro, hasta donde determinen puntualmente las disposiciones del mando para cada uno de los sistemas de armas estratégicas.

Este conflicto se ve especialmente agudizado en la práctica por una generalizada descentralización del control, necesaria por otro lado en atención al enorme poder destructivo de las armas nucleares. No se puede encerrar un moderno arsenal estratégico en un solo emplazamiento, ni siquiera en un pequeño número de ellos, sin que se corra el riesgo de que esos puntos sean identificados y destruidos por sorpresa, con la consiguiente incapacitación de toda la fuerza. Conscientes de esa realidad, los Estados Unidos han dispersado su capacidad nuclear en numerosos emplazamientos, algunos de ellos móviles, y entre centenares de oficiales. A lo que parece, la Unión Soviética ha hecho otro tanto.

Norteamérica asegura el control negativo a través de múltiples cortapisas físicas y limitaciones a los procedimientos organizativos, que hacen imposible que un individuo dispare armas nucleares e improbable que se congregue la necesaria combinación de personas –que van de dos a diez, o aún a más, según las circunstancias– en ausencia de la debida autoridad política. Con lo que impone este control disperso, el cronómetro se convierte en un factor importante en las operaciones con armas nucleares. En general, es bastante difícil mantener mecanismos de control negativo y conseguir al propio tiempo una precisa secuencia temporal de las operaciones ofensivas.

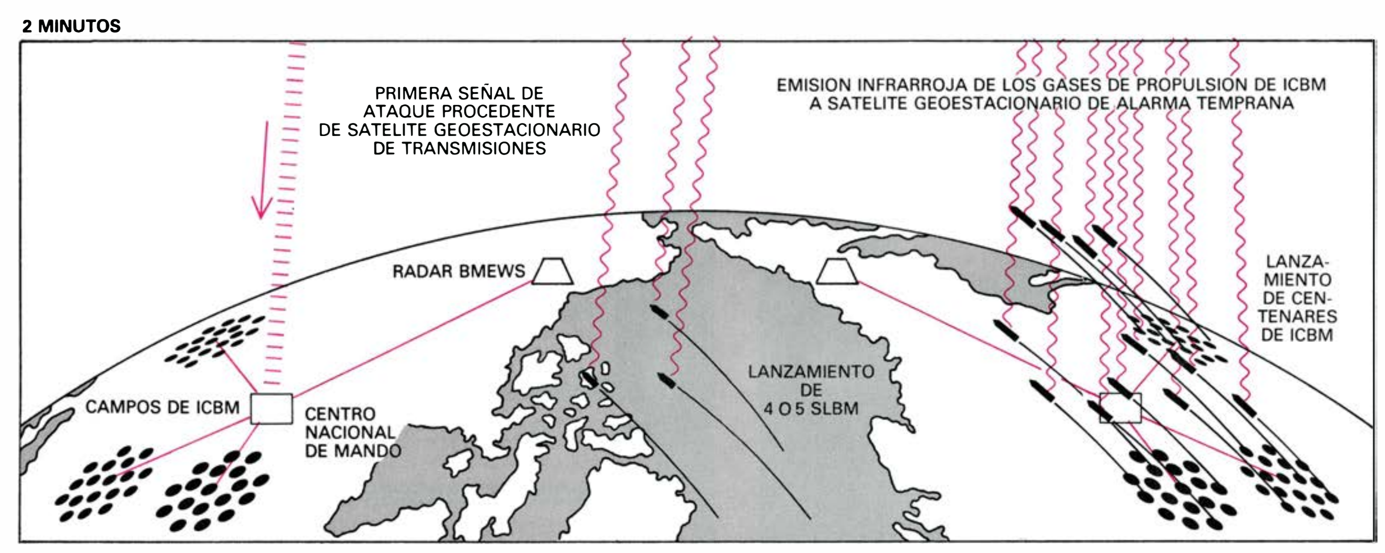
En un ataque contra silos se hace necesario un riguroso control del factor tiempo, a fin de reducir las interferencias entre las ojivas que explotan en la misma zona. En el caso de una base estadounidense típica de Minuteman, el ataque a los distintos silos tiene que efectuarse de sur a norte, y la cadencia de las ojivas que atacan a un mismo silo debe controlarse cuidadosamente en relación con la potencia y la precisión de las mismas. Es más, la acción debe cumplirse al primer intento, y en ella entran en juego de 2000 a 3000 ojivas. Los errores en el cálculo de los tiempos situarían los resultados del ataque muy por debajo de su potencial teórico.

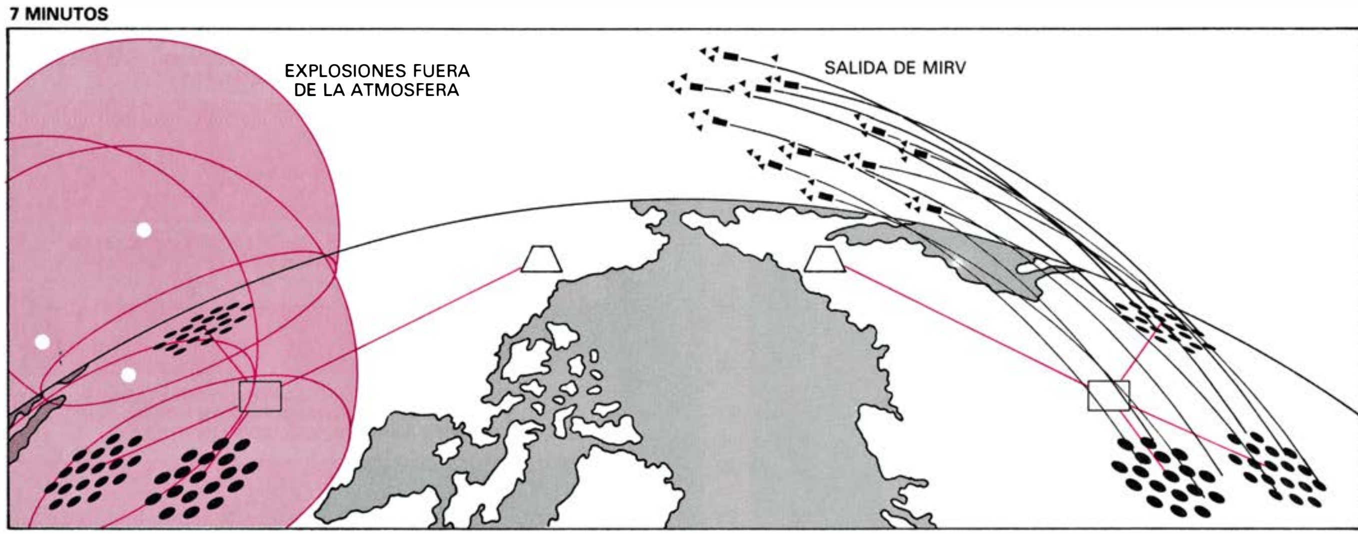
Cualquier agresor en potencia, capaz de combinar complejos cálculos técnicos con un poco de sentido común, apreciará que estas incertidumbres le quitan muchos de los estímulos que teóricamente le ofrece el lanzamiento de un ataque por sorpresa [véase "Vul-

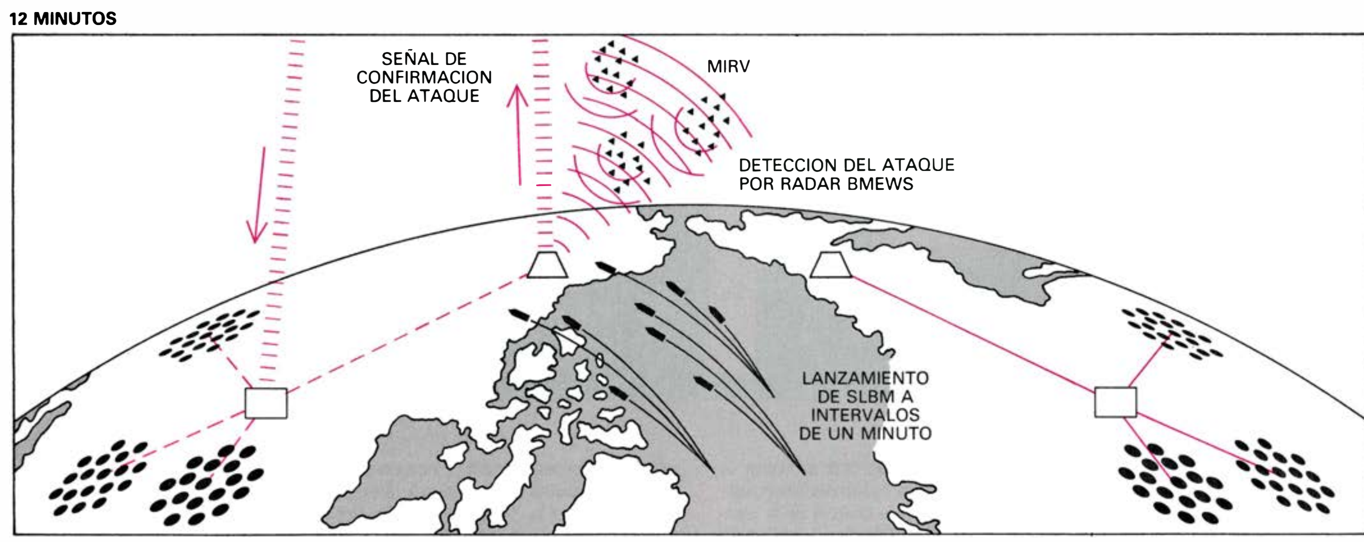


1. HIPOTETICA SUCESION DE ACONTECIMIENTOS tras la orden de ataque por sorpresa contra Estados Unidos por misiles balísticos intercontinentales (ICBM) soviéticos. Es imprescindible un estricto control de la sucesión temporal del proceso en un ataque contra silos de ICBM, así como para oponer unas medidas eficaces de respuesta precoz de contragolpe. El actual sistema norteamericano Minuteman quedaría expuesto a una quiebra de las comunicaciones durante el proceso de confirmación del ataque y de ejecución del contrataque. Una variante altamente automatizada, postulada por Richard L. Garwin, trataría de evitar dicha indefensión haciendo que los misiles se dispararan en la acción de represalia antes de transcurrir cinco minutos desde la primera señal de la orden de ataque. (SLBM son las siglas de *Sea-Launched Ballistic Missiles*, misiles balísticos instalados en submarinos.)

nerabilidad ante un ataque nuclear por sorpresa", por Matthew Bunn y Kosta Tsipis; INVESTIGACIÓN Y CIENCIA, enero de 1984]. Sí puede concebirse una operación a sangre fría, que explotara la ventaja que da la iniciativa para pasar por alto todos los mecanismos de control negativo de la parte atacante (sin dar ninguna señal externa) y que automatizase completamente las secuencias del ataque, resolviendo así el problema de sincronización cronológica. Ante esta posibilidad, teóricamente factible, los Estados Unidos no se han detenido en estudiar los problemas prácticos del atacante, y todos los esfuerzos se han centrado en buscar soluciones que aliviaran la vulnerabilidad de los silos dentro de los límites de esta definición técnica, en especial, aquella solución que no esté supeditada a errores de planificación cronológica en las secuencias del ataque. A juzgar por los diversos costos previstos para el programa MX,





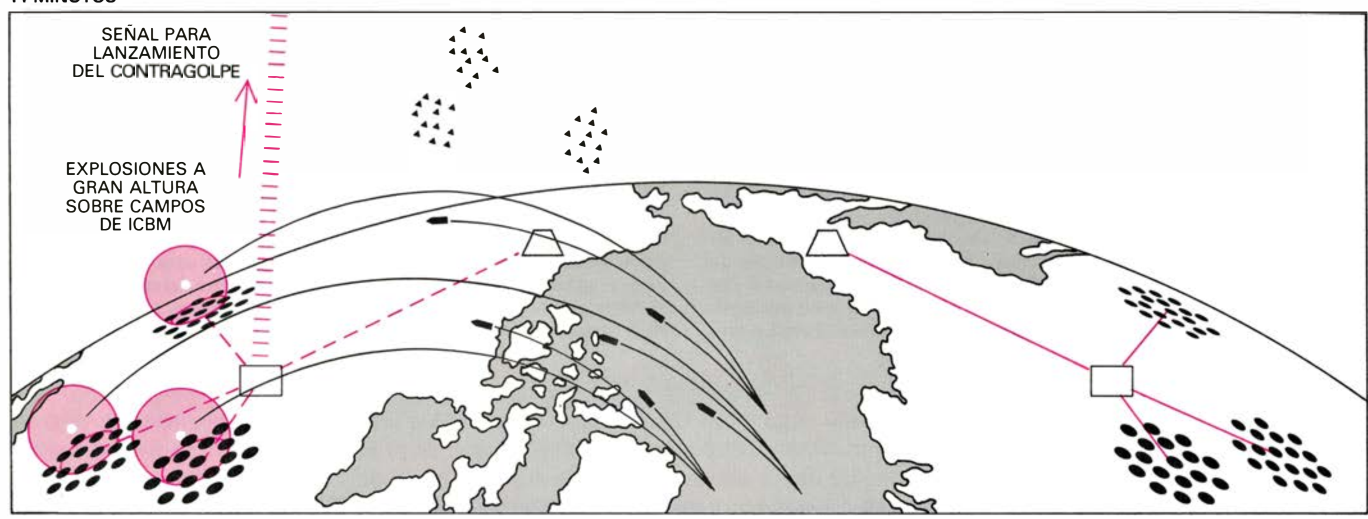


2. SEIS ETAPAS de la secuencia del hipotético ataque de la figura 1. En este supuesto, la decisión de emprender una acción de represalia no se tomaría hasta después de que el Sistema de Alerta Temprana de Misiles Balísticos (BMEWS, *Ballistic Missile Early Warning System*) confirmara el ataque. Los

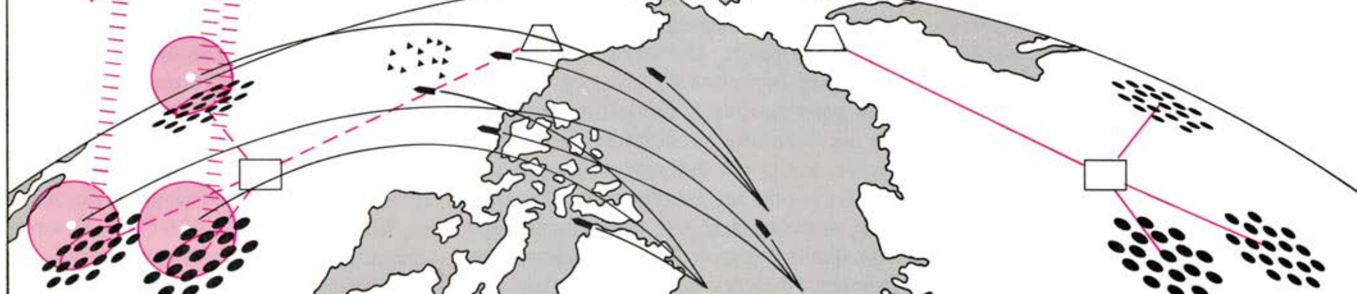
se han considerado aceptables gastos del orden de entre 30.000 y 50.000 millones de dólares.

También es necesario controlar la sucesión cronológica para llevar a cabo con éxito una política de respuesta precoz de contragolpe. Durante los primeros tres o cuatro minutos de su vuelo, los ICBM estadounidenses serían mucho más vulnerables a las explosiones nucleares producidas sobre sus bases que si continuasen encerrados en sus silos. Por ello, los errores de tiempo en la ejecución del lanzamiento ante ataque podrían favorecer la destrucción de los misiles norteamericanos, que no colaborar en su superviviencia. A escasa altura, en los primeros instantes de su fase de lanzamiento, estos misiles se verían expuestos a los escombros y partículas y a los fuertes vientos levantados por las explosiones a ras de suelo, cuya finalidad asignada sería destruir los silos. Fuera de la atmósfera, los misiles

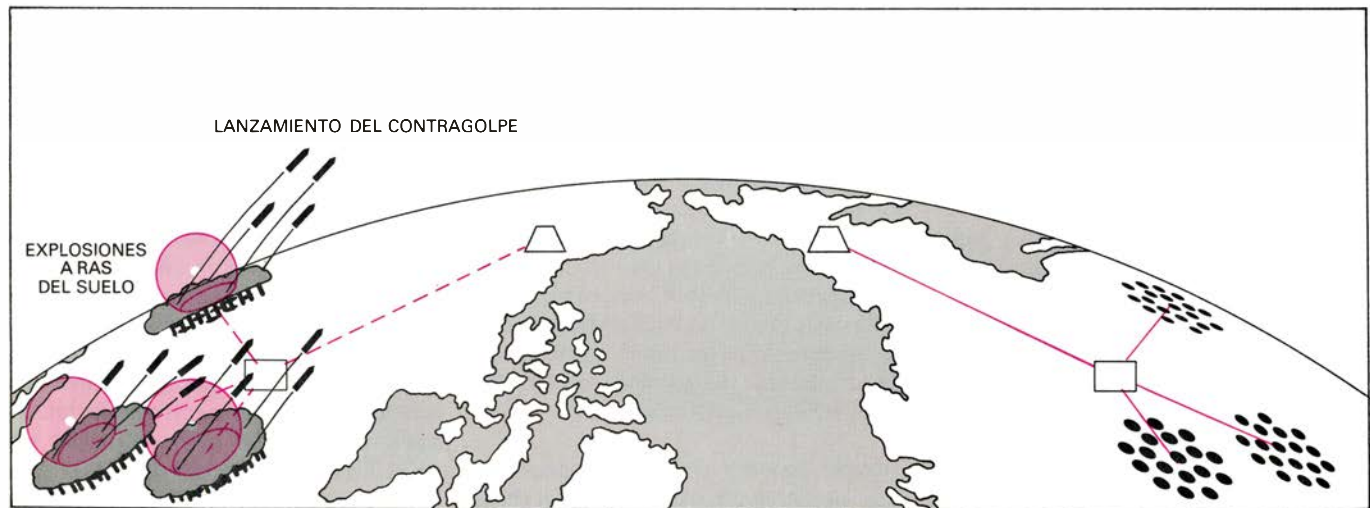
14 MINUTOS



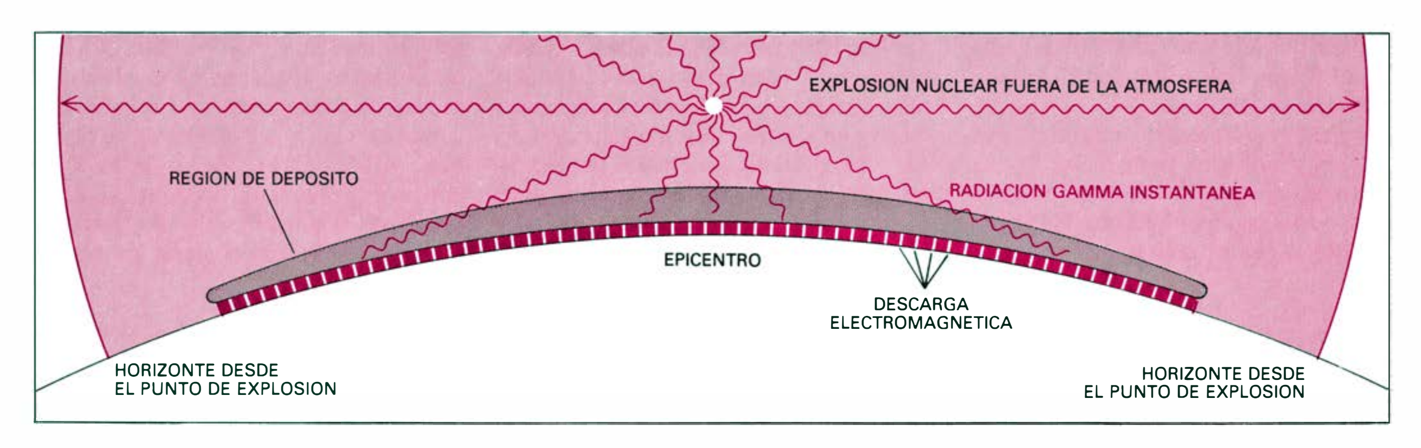
18 MINUTOS



25 MINUTOS



ICBM norteamericanos se lanzarían con tiempo suficiente para evitar los daños subsiguientes a las explosiones detonadas a ras de suelo; deberían atravesar, sin embargo, una barrera de explosiones provocadas a gran altura sobre todas las bases de ICBM para cerrar el paso a los misiles en su huida.



3. DESCARGA ELECTROMAGNETICA producida por una explosión nuclear de elevada potencia detonada a cerca de 500 kilómetros sobre la superficie de la tierra. Tendría efectos devastadores sobre las comunicaciones en una extensa área. Los rayos gamma instantáneos, emitidos por tal explosión, arrancarían electrones de las moléculas de aire de los altos niveles de la atmósfera en una zona circular, en forma de tarta, denominada "región de depósito". Los electrones libres experimentarían una aceleración radial por efecto del campo magnético terrestre, que los separaría de los iones más pesados y de carga positiva, produciéndose, por un complejo mecanismo, las condiciones necesarias para una descarga electromagnética de alto voltaje dirigida hacia abajo, que se manifestaría en forma de afloramientos eléctricos simultáneos en los conductores expuestos en tierra. En una explosión de la suficiente magnitud, los efectos destructores de esta descarga electromagnética de alto voltaje llegarían hasta el horizonte, la tangente a la tierra desde el punto de explosión (en este caso, hasta un radio de unos 2400 kilómetros desde el epicentro de la explosión). A su vez, la ruptura del campo magnético terrestre provocada por la explosión desencadenaría una segunda descarga electromagnética, de mucha menor magnitud, pero de mayor duración.

se verían expuestos al efecto perjudicial de los rayos X emanados de explosiones nucleares detonadas a gran altura. Puesto que la duración del vuelo de los ICBM soviéticos se calcula en 25 o 30 minutos, considerando que recoger, procesar y transmitir las primeras señales indicadoras del ataque (las estelas dejadas por los misiles atacantes en su fase de lanzamiento) tarda dos minutos, se contaría con algo menos de media hora para evaluar el ataque, decidir el contrataque y efectuar un lanzamiento de represalia, en respuesta a un ataque que concerniese únicamente a los ICBM. El tiempo con que se cuenta podría ser aún menor en caso de que los soviéticos utilizasen misiles lanzados desde submarinos para generar rayos X por medio de explosiones nucleares detonadas a una altitud de unos 160 kilómetros sobre los campos de ICBM.

Sin tener que ampliar los objetivos elegidos para un ataque directo, la Unión Soviética podría causar incluso serios daños al sistema de mando norteamericano, que cubre las necesarias funciones de evaluar el ataque, decidir las represalias y dar la orden de disparo. Los misiles lanzados desde submarinos soviéticos podrían detonarse, siete minutos después y a una altura de 480 kilómetros sobre el territorio continental de Estados Unidos, provocando explosiones nucleares que no causarían en tierra efectos explosivos, ni térmicos, ni de radiación inmediata, pero que crearían descargas electromagnéticas de dos tipos. Se generaría una primera descarga, del orden de 50.000 volt por metro, que apenas duraría una milésima de segundo, cuando los rayos gamma originados por las explosiones arrancaran electrones de las moléculas de aire de los altos niveles de la atmósfera y los electrones libres se dispusieran en círculo en torno a las líneas de fuerza magnéticas de la tierra. La descarga electromagnética resultante se recogería, según el ángulo de incidencia, en cualquier conductor expuesto y alcanzaría los equipos eléctricos o electrónicos conectados al conductor, interrumpiendo su funcionamiento o provocando su destrucción total.

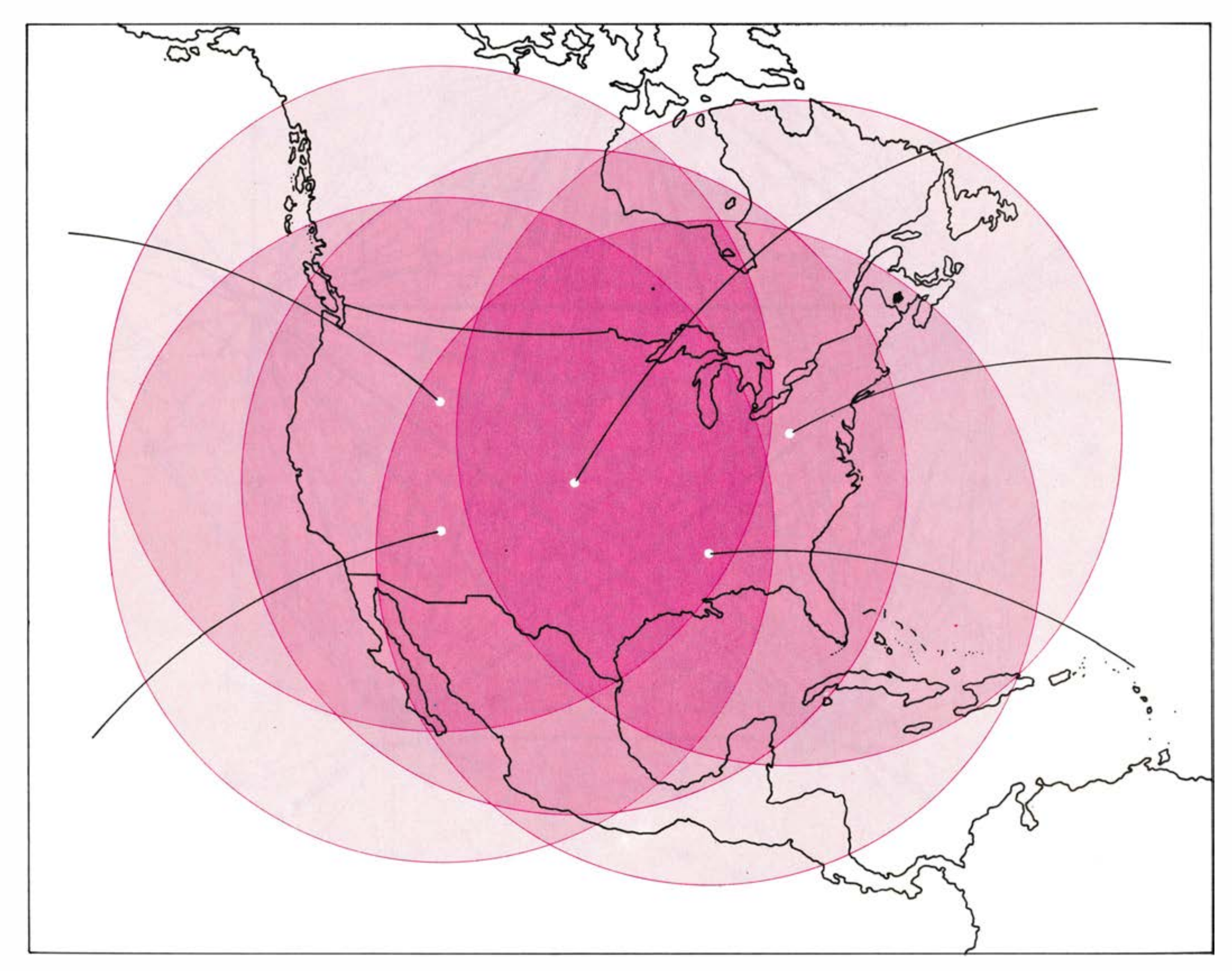
Una segunda descarga, de menor magnitud (decenas de volt por kilómetro) pero de mayor duración (centenares de segundos) resultaría de la alteración temporal del campo magnético terrestre por las explosiones. En este caso se produciría una importante ruptura de las líneas de comunicaciones terrestres en decenas de kilómetros. La complejidad de estos fenómenos y la imposibilidad de ensayarlos a escala real impiden la elaboración de cálculos precisos, pero está bien claro que los actuales sistemas de comunicaciones de líneas terrestres y las redes eléctricas locales correrían serios riesgos. No sabemos qué ocurriría exactamente si el territorio de Estados Unidos se viese sujeto en su totalidad a la caída de miles, o tal vez de decenas de miles, de relámpagos altamente acelerados, pero todo cálculo práctico debe predecir forzosamente fallos de los sistemas de comunicaciones, del suministro eléctrico y del equipo electrónico.

Cabe suponer que, en ausencia de descarga electromagnética u otros factores que provoquen deliberadamente la ruptura del sistema de mando, los sensores de inteligencia, los centros de proceso de la información y los sistemas de comunicaciones existentes en Estados Unidos podrían por sí solos cubrir las necesidades de un lanzamiento ante ataque contra el procedente de ICBM. Los sensores infrarrojos instalados en satélites geosincrónicos podrían detectar en cuestión de segundos el lanzamiento de un misil soviético y en pocos minutos podrían llegar a Estados Unidos las señales de alerta procesadas, indicando la magnitud aproximada del ataque y confirmando si éste va dirigido directamente contra Norteamérica. En 15 minutos, los radares del Sistema de Alerta Precoz de Misiles Balísticos (*Ballistic Missile Early Warning System*, BMEWS) instalados en Alaska, Groenlandia y Gran Bretaña ofrecerían una confirmación independiente e indicaciones generalizadas de las zonas de Estados Unidos objeto del ataque. A esta información habría que añadir la ofrecida a los 20 minutos por el radar antimisiles balísticos (ABM) instalado en Grand Forks, Dakota del Norte. Combinando estos indicadores se podría distinguir con seguridad un ataque real de cualquier otro fenómeno. A los pocos minutos de las primeras señales de alerta podrían celebrar el presidente y la cúpula militar una reunión por medio de sistemas de telecomunicación. Si con la primera confirmación del ataque se comunicase la autorización para responder, entrarían en acción todos los mecanismos de lanzamiento de misiles que comportan un control negativo dentro del plazo previsto para que los ICBM norteamericanos saliesen de sus silos hacia zonas de seguridad.

Sin embargo, una descarga electromagnética haría muy improbable el desarrollo pleno de esta secuencia. Los probables fallos de las líneas terrestres interrumpirían la transmisión de la información de alerta, pudiéndose producir la recepción de órdenes militares de alta velocidad de transmisión, bien al comienzo del ataque, bien después de los primeros indicadores de alerta, que quedaran sin ejecutar. Las comunicaciones de emergencia incólumes habrían de aprovechar en gran medida las transmisiones por radio a muy alta frecuencia entre aeronaves, que a su vez deberían despegar en respuesta a una alerta de ataque. (Pero, hoy por hoy, tales aeronaves no están suficientemente protegidas contra los efectos de una descarga electromagnética de alta intensidad.) La transmisión de mensajes a todo el país a través de esa red requeriría unos seis minutos. Con la ruptura de la red de transmisiones de alta frecuencia y los bloqueos subsiguientes, el sistema de comunicaciones que quedase en funcionamiento sólo podría operar a velocidades lentas, en las bandas de baja y muy baja frecuencia, y se necesitarían posiblemente casi 10 minutos para transmitir el mensaje en clave autorizando la acción de represalia.

Ante la distribución parcial de señales de alerta y la interrupción de las comunicaciones, los mecanismos normales de mando, sujetos a las exigencias del control negativo, retrasarían con toda seguridad el lanzamiento de salvaguardia de los misiles norteamericanos; la misma improvisación de medidas para tratar de agilizar el proceso acarrearía el riesgo de comportar resultados nefastos: lanzamientos con ligeros errores temporales expondrían a los misiles estadounidenses a su destrucción en la fase de lanzamiento. Dada la capacidad militar actual, parece confirmarse el criterio dominante según el cual una respuesta precoz de contragolpe, por insegura, resulta inviable.

De acuerdo con las medidas previstas por el actual plan quinquenal de defensa, se corregirán y se reforzarán puntos débiles del sistema de mando estadounidense. Las medidas más importantes conciernen a la protección, contra los efectos de las descargas electromagnéticas generadas por explosiones detonadas fuera de la atmósfera, de los principales puestos de mando en tierra y aire, de las aeronaves encargadas de la transmisión de comunicaciones y, en menor cuantía, de las principales líneas terrestres de comunicaciones. Una red de emisoras en tierra proporcionará información de la alerta transmitiendo en baja frecuencia, un enlace que no depende de cables, vulnerables por su propia naturaleza. Se crearán pequeñas estaciones móviles encargadas de recibir información pro-

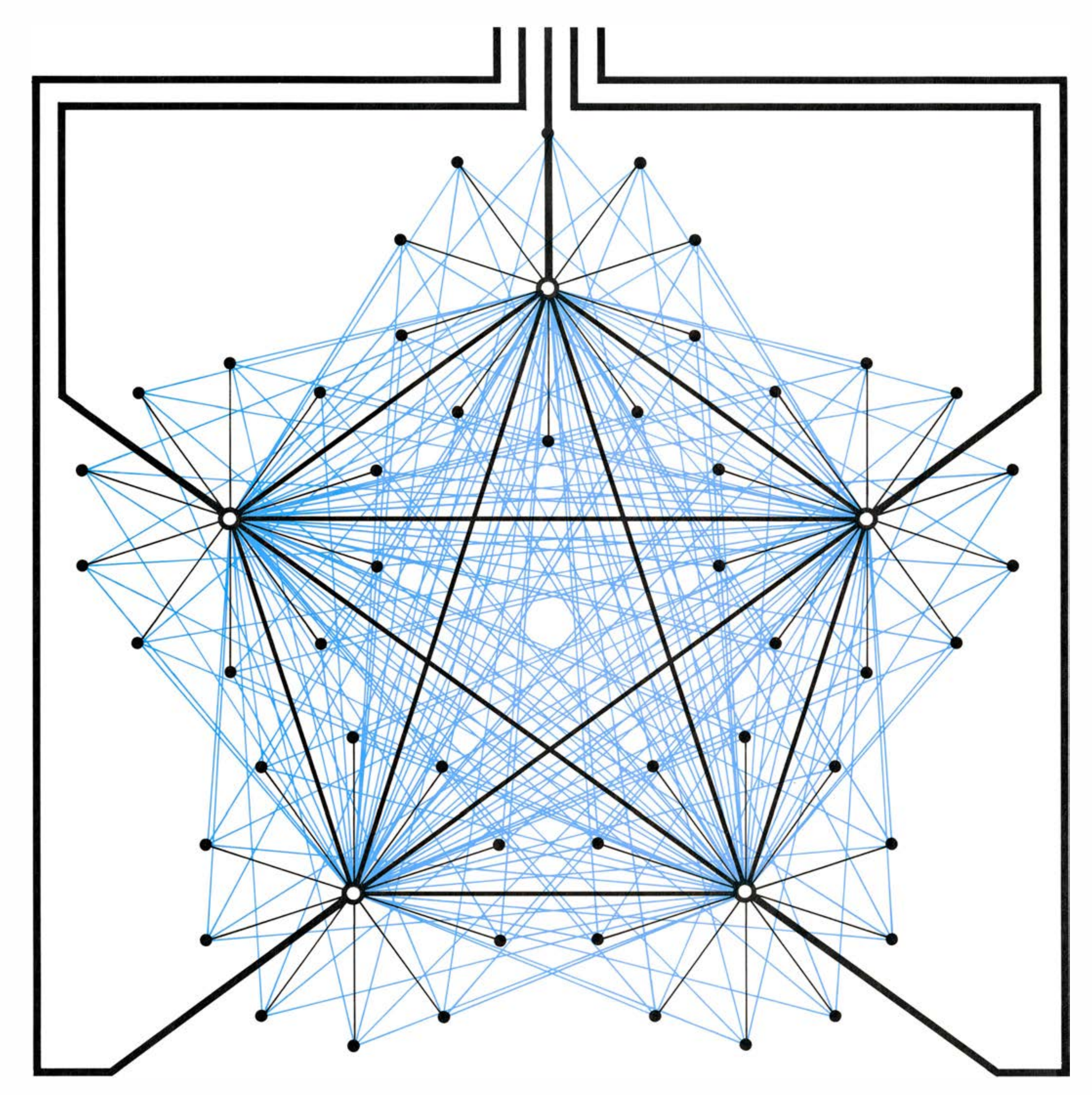


4. LA TOTAL COBERTURA DE ESTADOS UNIDOS por los efectos electromagnéticos destructivos de una explosión nuclear fuera de la atmósfera podría lograrse, en principio, detonando una sola ojiva nuclear de un megatón a una altura de cerca de 500 kilómetros sobre el centro del país. Dado que el daño causado por la descarga electromagnética de alto voltaje dependería en parte de la orientación de los conductores expuestos a la dirección del impulso, el hipotético atacante optaría probablemente por crear un modelo envolvente de efectos, confiando cuatro o cinco ojivas nucleares a esta fase del ataque. Si las cargas se lanzaran desde submarinos situados en la zona donde suelen patrullar, alcanzarían sus puntos de explosión en unos siete minutos.

cedente de satélites de alerta precoz y se dotará a los bombarderos de receptores de muy baja frecuencia, lo que permitirá fortalecer notablemente las comunicaciones por radio, más allá de la línea de percepción de la red de aeronaves de enlace.

Estas mejoras dificultarán la quiebra del sistema de mando norteamericano, sin que para ello tengan que ampliarse, de modo importante y provocador, los objetivos a atacar directamente con los efectos normalmente previstos de la onda expansiva de las explosiones nucleares. Pero el programa no se ha elaborado para proporcionar esa capacidad coherente y fiable de respuesta precoz de contragolpe que se ha descrito de modo teórico; en la práctica, dejará a Estados Unidos muy por debajo de esa capacidad, diferencia ésta que puede medirse en décadas de esfuerzos. Antes de que pueda alcanzarse el consenso necesario para emprender esos esfuerzos, habrá que plantearse seriamente cómo inciden en la concepción global del problema el estado del sistema de mando y los aspectos de control.

Desde esta última perspectiva, más general, la reconocida inviabilidad de una política de lanzamiento ante ataque plantea problemas de seguridad mucho más graves que el de la vulnerabilidad de los silos. Si el sistema de mando norteamericano puede resentirse de modo importante por los efectos electromagnéticos producidos por un pequeño número de explosiones fuera

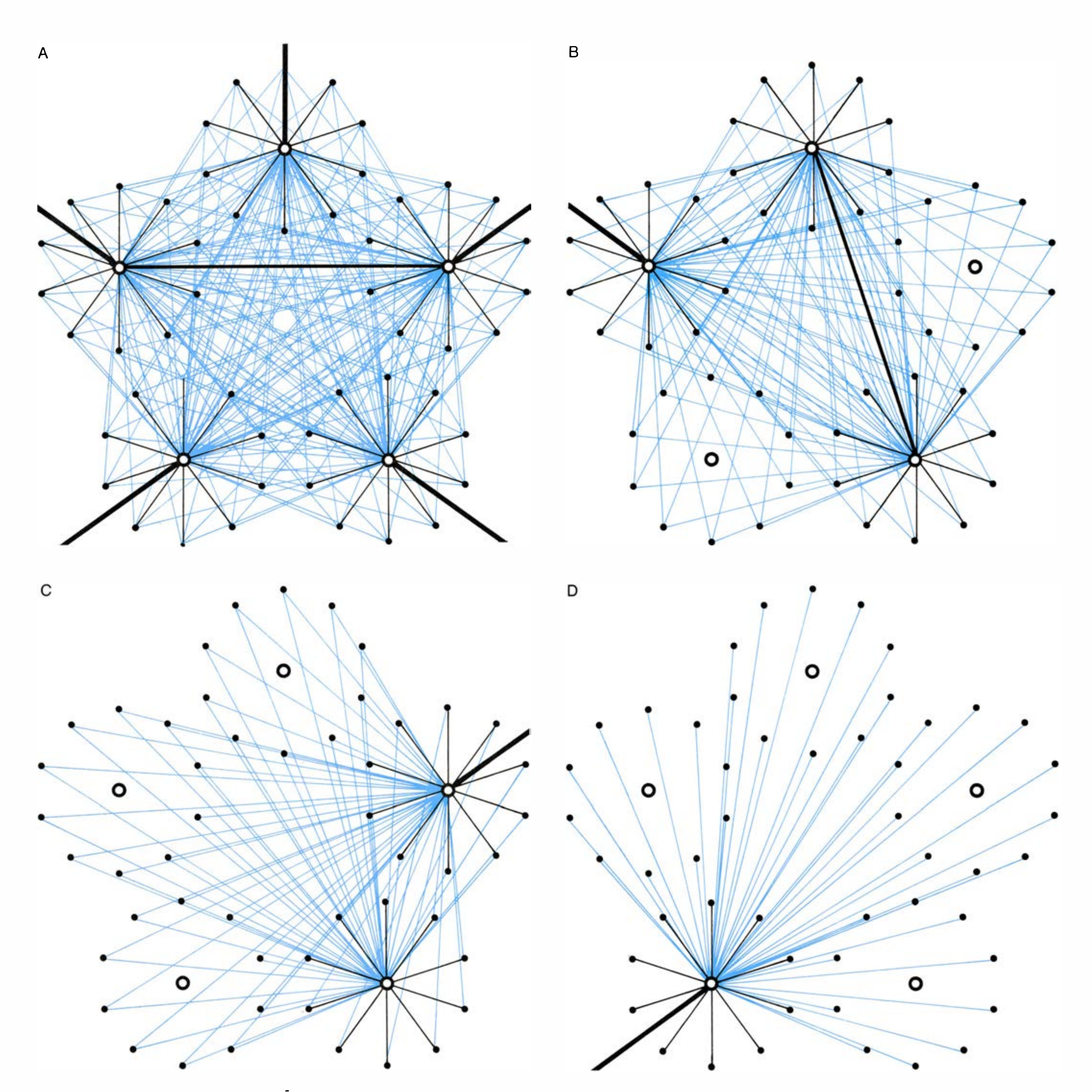


5. SISTEMA DE MANDO de los 50 misiles de un escuadrón de Minuteman II, esquematizado. El diseño se funda en un modelo simplificado del que desarrollaron por ordenador Paul Morawski y Bruce G. Blair, de la Institución Brookings. Las líneas de comunicaciones unen los cinco centros de control con el mundo exterior (*líneas negras gruesas*), cada centro con cada uno de los restantes (*líneas negras de trazo intermedio*) y con cada uno de los 50 misiles. Cada centro suele tener responsabilidad de primer grado sobre 10 misiles (*líneas negras de trazo fino*) y responsabilidades de supervisión sobre los 40 restantes (*líneas de color*). El programa de ordenador incorpora complejas reglas que contienen procedimientos de control positivo y negativo.

de la atmósfera, lo destruirán los efectos directos de la onda expansiva de algunos centenares de armas, o quedará seriamente dañado con la detonación de sólo 50 armas. Estos ataques no requerirían especiales exigencias operativas, en términos de comportamiento técnico de las ojivas ni del grado de posibles interferencias entre ellas. Si el lanzamiento de salvaguardia de los misiles durante un ataque resulta amenazado por la quiebra del sistema de mando, cabe poner en tela de juicio el posterior disparo de las armas que sobrevivan al ataque, pudiendo albergarse también serias dudas en cuanto a la ejecución de acciones de represalia planeadas coherentemente.

Además, aunque se cree que el sistema de mando soviético está más protegido, los efectos destructores de las armas nucleares alcanzan de por sí tal magnitud para cualquier red de mando que las diferencias de exposición entre uno y otro bando antagónico no parecen revestir mayor interés, habida cuenta de la potencia de fuego de que podría hacerse uso. Por tanto, cabe concluir que para ambos bloques la vulnerabilidad del conjunto de sus sistemas de mando estratégico encierra una importancia mucho mayor que la vulnerabilidad de sus armas; y ello es así desde que hace dos décadas comenzaron, Este y Oeste, el primer despliegue a gran escala de misiles balísticos ofensivos.

Comparada con la vulnerabilidad de los silos, la vulnerabilidad del sistema



6. CUATRO SITUACIONES DE DAÑO en el sistema de mando de un escuadrón de Minuteman II, expuestas por orden creciente de gravedad. En la situación *A* se asignan dos centros de control al estado operativo calificado de normal (con todos los enlaces de comunicaciones intactos) y tres se hallan en situación vertical (han perdido los enlaces con los otros centros). En la situación *B*, un centro es vertical, dos son horizontales (han perdido los enlaces con los comandantes del exterior) y dos han desaparecido (todos los enlaces destruidos). En la situación *C*, un centro es vertical, otro mudo (sólo le quedan los enlaces con los misiles) y tres centros han desaparecido. En la situación *D* un centro es vertical, mientras que los otros cuatro han desaparecido.

de mando ofrece un aliciente mucho más poderoso para iniciar un ataque antes de que se sufran los daños, aliciente que no procede de la oportunidad definida teóricamente por complejos cálculos teóricos, sino de temores de orden práctico ante una decisiva derrota en una guerra que no pueda evitarse. Sea cual fuere la política de seguridad nacional declarada en tiempos de paz, este aliciente ejerce sobre los responsables militares, tanto en Estados Unidos como en la Unión Soviética, presiones potencialmente abrumadoras en favor de decidir una acción por sorpresa de primer golpe en caso de grave crisis, cuando tengan que enfrentarse con las perspectivas de una guerra inevitable.

Para evaluar las consecuencias convendría contar con cálculos de los efectos de un ataque nuclear sobre todo el sistema de mando, a imagen de los que se usan para evaluar los efectos sobre silos de misiles. Tales cálculos no se han debatido en público, por la natural reserva de los asuntos relacionados con la seguridad y porque el problema conlleva complejas interacciones de organización técnica y humana, y no se ha efectuado una rigurosa evaluación de esos problemas. A pesar de lo cual, y sin pretender una evaluación estadísticamente válida, podemos esbozar un modelo conceptual de las interacciones entre las vulnerabilidades del sistema de mando, los métodos para asegurar los controles positivo y negativo y las secuencias temporales del ataque y la respuesta al mismo.

Paul Morawsky y Bruce G. Blair, de la Institución Brookings, han elaborado un modelo simple, bajo la forma de programa de ordenador, que incorpora las disposiciones para el mando y las normas de procedimiento para el lanzamiento de los 50 misiles de un escuadrón Minuteman II publicadas a mediados de la década de 1970. Este sistema constituye una parte muy pequeña de la red de mando estratégico de Estados Unidos e incorpora sólo una fracción de su complejidad y algo, muy poco, de su profundidad jerárquica. Refleja solamente parte de las operaciones de un auténtico escuadrón. Sin embargo, en esta representación tan simplificada, ciertos resultados parecen apuntar algunas de las presiones internas con las que se enfrentarían, presumiblemente, las fuerzas estratégicas reales.

El modelo muestra cinco centros de control distintos y enlazados entre sí y con cada uno de los 50 misiles [*véase la figura 5*]. Las normas que determinan las acciones de cada centro reflejan los procedimientos de control negativo y positivo que deben seguirse para activar los misiles Minuteman II. Cada centro tiene una responsabilidad de primer grado sobre 10 misiles (un ejemplo de control positivo) y responsabilidades de control sobre los 40 restantes (control negativo). Asimismo, cada centro es sustituto de otro, de modo que, si se destruyera uno, su sustituto asumiría la responsabilidad de los misiles (control positivo) que le correspondían. Cada centro ejecutará lo dispuesto para establecer un plan conjunto de objetivos para los 50 misiles, asignar objetivos a los misiles que se hallen bajo su jurisdicción y lanzar o retener en tierra sus misiles primarios (control positivo), pero sólo tras cumplir escrupulosamente con lo dispuesto para corroborar que las instrucciones son válidas (control negativo).

Para cada vértice de este pentágono, el procedimiento de convalidación exige la confirmación de otro centro de control del sistema (control negativo). Sólo se lanzará un misil hacia el objetivo que tiene asignado si la orden recibida en tal sentido la confirma un segundo centro de control. Este segundo vetará las órdenes de lanzamiento del primero si las considera inadecuadas: porque no haya recibido ni convalidado una orden de autorización (control negativo) o porque no haya satisfecho sus

	PLAN I: REPRESALIA CONVENCIONAL		PLAN II: REPRESALIA CON URGENCIA DE TIEMPO		PLAN III: REPRESALIA LIMITADA	
SITUACIONES DE LOS CENTROS DE MANDO	SE RECIBEN LAS ORDENES ANTES QUE EL DAÑO	SE RECIBEN LAS ORDENES DESPUES DEL DAÑO	SE RECIBEN LAS ORDENES ANTES QUE EL DAÑO	SE RECIBEN LAS ORDENES DESPUES DEL DAÑO	SE RECIBEN LAS ORDENES ANTES QUE EL DAÑO	SE RECIBEN LAS ORDENES DESPUES DEL DAÑO
A	P = 100 N = 0	P = 94 N = 1	P = 100 N = 0	P = 93 N = 0	P = 100 N = 0	P = 95 N = 3
B	P = 99 N = 0	P = 0 N = 0	P = 95 N = 3	P = 0 N = 0	P = 96 N = 8	P = 0 N = 0
C	P = 98 N = 2	P = 0 N = 0	P = 91 N = 5	P = 0 N = 0	P = 90 N = 14	P = 0 N = 0
D	P = 42 N = 1	P = 44 N = 2	P = 19 N = 1	P = 23 N = 0	P = 36 N = 0	P = 40 N = 0

7. RESULTADOS OBTENIDOS A PARTIR DEL MODELO de operación de un escuadrón de Minuteman II. Se indica, por separado, la puntuación media de los controles positivo y negativo tras 10 iteraciones del modelo. La puntuación del control positivo (*P*) da el porcentaje de acciones de contragolpe ejecutadas con éxito, definidas asignando valores ponderados a cuatro categorías de objetivos según la relación 1:2:3:4. Las categorías 2 y 4 representan objetivos militares "urgentes"; su valor se reduce, respectivamente, en un 50 por ciento a los 90 minutos y en un 80 por ciento a los 60 minutos. Se juzga perfecta (*P* = 100) la operación en que todos los misiles se lanzan hacia sus objetivos en el tiempo previsto. La puntuación del control negativo (*N*) ofrece el porcentaje de misiles lanzados hacia objetivos no asignados a ellos con respecto al número total de misiles disparados. En el plan I se asignan los misiles de forma bastante paritaria a las cuatro categorías de objetivos. En el plan II se asignan 15 misiles a la categoría 3 y 35 a la categoría 4. El plan III contempla el lanzamiento de 10 misiles a la categoría 3, y 10 a la cuatro, dejando en reserva 30 misiles. Uno de los problemas detectados es que, en condiciones intermedias de daño (*B*, *C*), se produce un fallo catastrófico del control positivo si las instrucciones de contragolpe se reciben después de sufrir los daños. El sistema logra efectuar el lanzamiento de contragolpe en la situación de peor daño (*D*), pero con un retraso de 90 minutos. En las situaciones *B* y *C* se produce asimismo un fallo, estadísticamente inferior, del control negativo; se disparan algunos misiles cuyo despegue debería haberse suspendido.

asignaciones de objetivos para todos los misiles (control positivo). Un veto cancelaría la orden de lanzamiento de cualquier centro (control negativo), pero no tendría efecto si dos centros coincidiesen en la orden de lanzamiento (control positivo). Si la orden de lanzamiento de un centro no recibiera confirmación ni veto, se ejecutaría (control positivo), pero con cierto retraso (control negativo). Estas normas constituyen sólo parte de la complejidad del procedimiento para el lanzamiento de misiles balísticos cuando la cadena de mando no está unificada.

	PLAN II SEGUIDO DE PLAN I		PLAN III SEGUIDO DE PLAN I	
PLANES DE DESIGNACION DE OBJETIVOS	LAS NUEVAS ORDENES LLEGAN ANTES QUE EL DAÑO	LAS NUEVAS ORDENES LLEGAN DESPUES DEL DAÑO	LAS NUEVAS ORDENES LLEGAN ANTES QUE EL DAÑO	LAS NUEVAS ORDENES LLEGAN DESPUES DEL DAÑO
PLANES CONVENCIONALES	P = 98 N = 2	P = 51 N = 13	P = 98 N = 2	P = 69 N = 3
REASIGNACION DE OBJETIVOS APLICADA AL PLAN I	P = 87 N = 7	P = 49 N = 15	P = 87 N = 7	P = 68 N = 13

8. NO ES FIABLE YA EL MODELO si se le exige el cumplimiento de órdenes más complejas. En este caso, el sistema de mando recibe instrucciones en el sentido de ejecutar los planes II o III y cambiar luego al plan I. Se sufre el daño en el intervalo comprendido entre la recepción de los dos bloques de instrucciones o tras la recepción de las últimas instrucciones. En general, los resultados reflejan fallos de los dos tipos de control, positivo y negativo. Cuando se modifican las asignaciones de objetivos añadidos al plan normal (*fila inferior*) el proceso entero tarda más en cumplirse y quedan pendientes algunas operaciones de reasignación de objetivos cuando llega el momento de lanzar los misiles. (Dibujos de Jerome Kuhl.)

El modelo ejecuta estos procedimientos en una secuencia cronológica continuada y eficaz y excluye los errores que cabe prever si hubiesen de llevarlos a cabo seres humanos desconcertados y atemorizados, al borde de la guerra nuclear. Pero tiene en cuenta los daños físicos que pueda causar el ataque del adversario sobre las líneas de comunicaciones, asignando cada centro a uno de los cinco estados operativos siguientes: normal (con todos los enlaces de comunicaciones intactos); vertical (con pérdida de los enlaces con otros centros); horizontal (con pérdida del enlace con los comandantes del exterior); mudo (sólo se mantienen los enlaces con los misiles); y desaparecido (ha quedado inutilizado, con todos los enlaces destruidos). De este abanico de posibilidades surgen hasta 525 condiciones de grave daño parcial en una red de cinco centros. Incluso excluyendo todos los errores de procedimiento y sin posibilidad alguna de iniciar ninguna acción fuera de las ordenadas por las instrucciones que se han dado, las interacciones entre la cronología de los daños sufridos por la red y la ejecución de los procedimientos de control establecidos podrían producir resultados distintos de los buscados por las instrucciones. Fallos que afectan al control negativo y al positivo.

Se han obtenido algunos resultados representativos del modelo bajo cuatro situaciones diferentes de daños en la red del centro de mando, ordenadas en cierta medida por su diferente grado de gravedad [*véase la figura* 7]. En todos lo casos, las instrucciones ordenan simplemente la ejecución de un solo plan de objetivos para todos los misiles. Se hicieron pruebas con tres diferentes planes de objetivos: dura represalia sin restricciones, castigo de objetivos militares y contragolpe circunscrito. En cada caso se compara el funcionamiento del sistema cuando se reciben las instrucciones poco antes de acusar el golpe con su puesta en acción cuando el daño precede a las instrucciones. A pesar del grave daño descrito, el modelo funciona bien en general, pero se advierten algunos problemas importantes. Las situaciones de daño medio (denominadas *B* y *C*) provocan un fallo catastrófico del control positivo cuando las instrucciones para decretar la represalia llegan después de acusar el golpe. En tales circunstancias, las reglas de procedimiento provocan contradicciones internas y no se lanza ningún misil. En la situación de haber sufrido el más grave daño (*D*), el sistema consigue replicar con una acción de represalia ordenada, pero con un retraso considerable (90 minutos), lo que reduce el supuesto valor del ataque con urgencia a objetivos como aeródromos, bases de submarinos, misiles móviles y centros de mando. El control negativo suele tener éxito; mas en las situaciones de daño medio se despliegan misiles que debían retenerse en tierra en la hipótesis o plan de contragolpe circunscrito.

El sistema modelizado de mando pierde eficacia en cuanto se le pide que ejecute instrucciones más complejas, que supongan dirigir el proceso de represalia bajo las presiones de tiempo que impondrían las operaciones de respuesta precoz de contragolpe [*véase la figura* 8]. Primero se dan instrucciones al sistema para que ejecute uno de los planes limitados y luego se cambia hacia una acción de represalia total. El daño se sufre en el intervalo de tiempo que transcurre entre la recepción de las dos series de instrucciones o después de la llegada de las nuevas instrucciones. El promedio de resultados obtenidos en las diez pruebas que se ilustran señala fallos de control positivo y negativo.

El análisis de los resultados de los distintos ensayos, considerados por separado, revela dos fuentes diferentes de fracaso. Al pasar de un plan de objetivos a otro, se mantuvo el primero en dos de las diez pruebas, a pesar de las instrucciones posteriores, según determinó el factor azar de cuál de los centros sería el primero en iniciar las operaciones. Cuando se imponían mayores exigencias (modificar el plan normal asignando objetivos suplementarios a determinados misiles) el proceso completo se ejecutaba en más tiempo. En este caso, las diez pruebas exprimentaron el cambio de planes ordenados, pero algunas de las operaciones de cambio de objetivos no se habían realizado plenamente cuando se ordenó el lanzamiento de los misiles. Este proceso de cambio de planes se parece, en sus exigencias, al de poner fin a un plan de lanzamiento antes de su plena ejecución. En situaciones en que se haya producido ya un cierto daño, no puede suspenderse la acción una vez que la ejecución de las instrucciones esté ya en marcha.

El modelo Brookings tiene más de ejercicio de lógica que de evaluación empírica. Además, sus resultados concretos no son tan importantes como sus implicaciones generales. En la práctica, los misiles Minuteman pueden controlarse desde un puesto de mando con sede en el aire o a través de una red terrestre de controles fijos. Si la aeronave pudiese evitar por sí misma una grave quiebra de las comunicaciones, se obviarían algunos fallos que se señalan entre los resultados del modelo. Pero, aun actuando con gran profundidad y acierto en el procedimiento, el sistema real de mando estratégico no puede evitar los problemas inherentes

que muestra el modelo. La ruptura parcial de las comunicaciones constituye la principal amenaza contra la coherencia de las disposiciones de control, que se hallan repartidas; lo que provoca un sinfín de situaciones posibles demasiado amplio para permitir una clara previsión y agudiza la pugna entre control positivo y negativo.

La cúpula militar consciente de estos problemas no se fiará del funcionamiento de sus fuerzas al sobrevenir una grave ruptura del sistema de mando, ni sabrá cómo conseguir el óptimo equilibrio entre el control negativo y el positivo. Este factor aviva la tendencia a iniciar operaciones ofensivas antes de que se produzca el daño, a fin de garantizar tanto el ajuste temporal como la cobertura de los blancos, necesarios para lograr los objetivos militares. Convierte al sistema de mando adversario en el blanco más deseado para sus propósitos y el medio que mejor se presenta de alcanzar la victoria. De este modo, las dos fuerzas estratégicas que se enfrentan parecen imponer la una a la otra poderosos alicientes para un ataque por sorpresa como el medio más seguro para dominar en un conflicto nuclear.

Esta conclusión presenta una amenaza de guerra muy diferente de las que normalmente se han venido conjugando. Durante varias décadas, la disuasión ha parecido suficientemente potente para garantizar que los dirigentes políticos eviten la guerra nuclear en tanto en cuanto los acontecimientos permaneciesen sujetos a sus decisiones conscientes. Los sistemas de mando existentes proporcionan una seguridad razonable de que un fallo técnico o un error humano no serán la causa desencadenante de una catástrofe. Por esa razón, en circunstancias normales se ha venido considerando la guerra nuclear como una posibilidad bastante remota. Sin embargo, las dos superpotencias nucleares están sujetas a tensiones peligrosísimas en un ambiente de crisis. Una y otra se ven comprometidas a conducir una guerra nuclear por medio de operaciones ofensivas meticulosamente programadas y cuidadosamente ajustadas en el tiempo, que exigen una exacta coordinación entre las diferentes armas que en ella se vean implicadas. Aunque, en principio, ambas aspiran a tener la capacidad de poder ajustar esas operaciones a las circunstancias concretas de la guerra, y aunque ambas han realizado considerables inversiones en la protección de sus fuerzas para la consecución de dicho fin, en realidad ninguna puede garantizar tal coordinación después de haber sufrido las primeras 50 o 100 explosiones nucleares. Si alguna vez la guerra pareciese inevitable, los jefes militares de cada parte a quienes corresponda la ejecución de las misiones que les han sido asignadas intentarán inevitablemente hacerse con el poder que les permita llevar la iniciativa, sea cual fuere la anterior política de seguridad nacional. Y lo harían con un rigor que dependería directamente de la intensidad de la crisis. En ese punto, sus presiones sobre los dirigentes políticos serían apremiantes. Aunque no hay razón para dudar de su permanente deseo de evitar la guerra, sí existen fuertes razones para dudar de su capacidad de poder contener a sus respectivos estados mayores.

Puesto que no ha habido aún ninguna crisis de suficiente gravedad que haya posibilitado calibrar esas presiones, ni las ideas estratégicas predominantes han definido con precisión el problema, éste no se ha apreciado en su magnitud. La crisis cubana de los misiles de 1962 tuvo lugar antes de que ninguna de las dos fuerzas estratégicas hubiese alcanzado la plena madurez; la crisis de Oriente Medio de 1973 no puso en estado de alerta las armas nucleares de ambos mandos. Además, los sistemas de armas que más agravan el problema no son los de avanzada precisión, cuyas capacidades ofensivas se conocen mejor analíticamente, sino los que afectan a la coordinación y al ajuste temporal de las operaciones estratégicas. Por ejemplo, el misil norteamericano de alcance medio Pershing II, que se está instalando en Europa, reduce considerablemente el tiempo de que dispondría el sistema de mando ruso para reaccionar. Igualmente inquietante es la capacidad soviética, presentada a menudo como una amenaza, aunque todavía sin definir claramente, para "colocar a Estados Unidos en situación semejante". Por añadidura, las armas contra satélites, capaces de atacar sistemas vitales de alarma y comunicaciones, reducirían gravemente la cohesión del sistema de mando adversario.

No se ha reconocido aún en el debate público sobre la política de seguridad nacional la importancia de esos avances armamentísticos. Por ello y a la postre, la idea de una respuesta precoz de contragolpe no ha de contemplarse como una solución teórica del problema de la vulnerabilidad de los silos, ni como una política poco práctica que pueda evitarse, sino como un síntoma de la situación cargada de peligros en que se halla la seguridad y que exige más atención que la que hasta ahora se le ha venido prestando.

Efectos atmosféricos de El Chichón

Esta erupción de 1982, relativamente pequeña y rica en azufre, inyectó en la estratosfera una fina bruma de gotículas de ácido sulfúrico, más densa que ninguna otra nube volcánica desde la gran erupción del Krakatoa en 1883

Michael R. Rampino y Stephen Self

La erupción del volcán mexicano de El Chichón, a últimos de marzo y primeros de abril de 1982, que no alcanzó proporciones extraordinarias, inyectó en la estratosfera una cantidad ingente de materia volcánica. Envió cenizas y gases volcánicos a más de 25 kilómetros de altura. Las observaciones por satélite, realizadas inmediatamente después, mostraron que la materia se movía en la estratosfera con rapidez hacia el Oeste; al cabo de pocas semanas, un fino velo de materia ceñía el globo. En menos de un año, la nube estratosférica había recubierto el hemisferio Norte entero y buena parte del Sur.

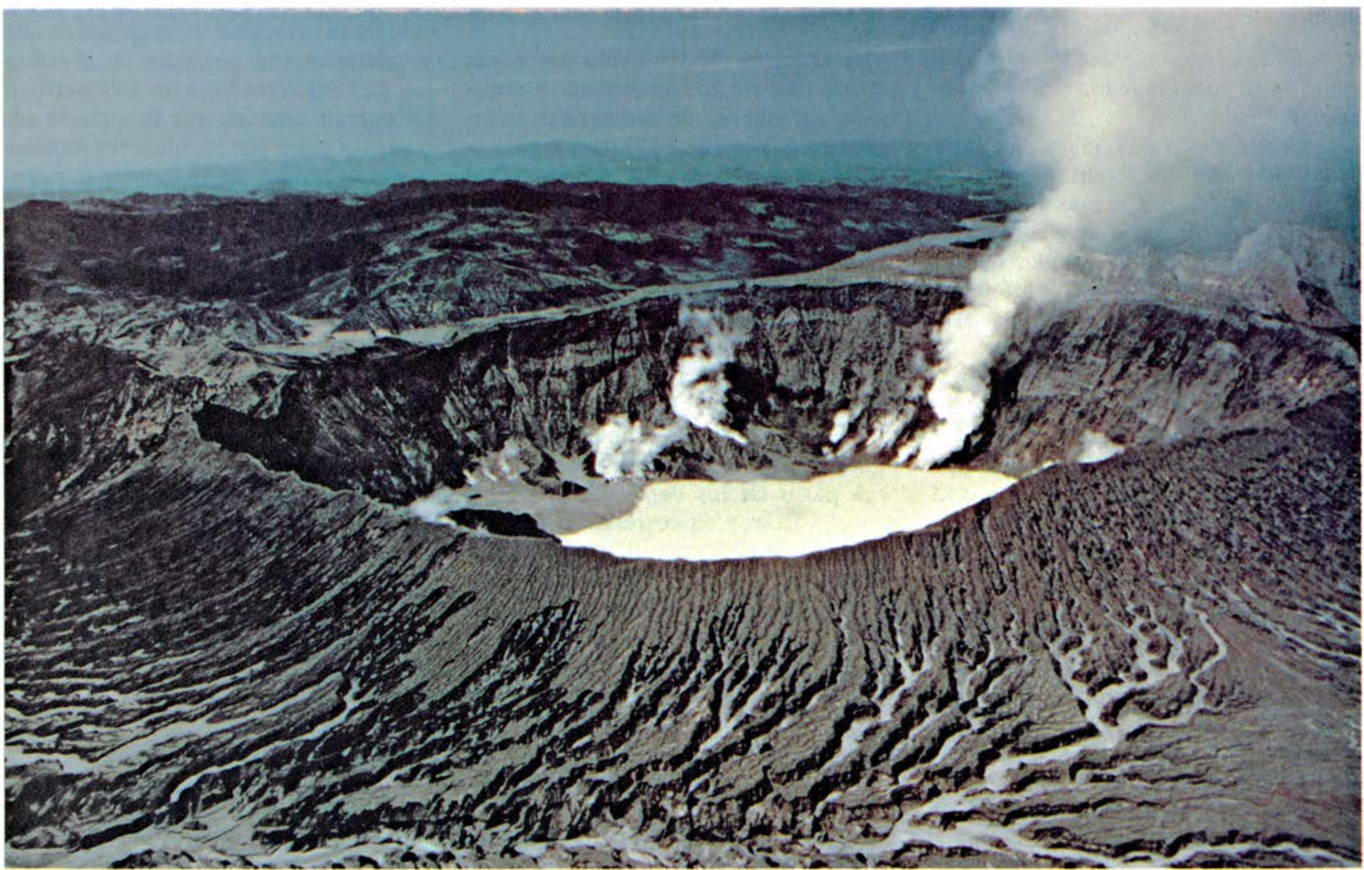
Desde hace algún tiempo se sabe que las nubes volcánicas de la estratosfera pueden afectar al clima mundial, al provocar, principalmente, un descenso de la temperatura media global o hemisférica. Se creía hasta hace muy poco tiempo que el volumen de cenizas finas y polvos expulsados en una erupción explosiva era una buena medida de la densidad de la nube producida, y por tanto, de su efecto sobre el clima. En este sentido, el índice de velo de polvo introducido en 1970 por el climatólogo británico Hubert H. Lamb clasifica las nubes creadas por erupciones históricas en razón de las estimaciones de la cantidad de ceniza expulsada en la erupción, datos de los efectos ópticos en la atmósfera y disminución de la temperatura en superficie en los años siguientes a la erupción. Sin embargo, a lo largo de los diez últimos años ha quedado claro que la mayor parte del polvo se deposita, en pocos meses, fuera de la estratosfera y que las nubes volcánicas de larga vida no están constituidas por polvo, sino por un aerosol de gotículas de ácido sulfúrico. Por consiguiente, la cantidad de gases ricos en azufre liberados en una erupción volcánica explosiva constituye un indicador de sus efectos atmosféricos mejor que el volumen de cenizas desprendido por la erupción.

La erupción de El Chichón demostró, por primera vez, que una erupción volcánica bastante pequeña, aunque rica en azufre, producía una nube estratosférica densa y extensa. Si bien El Chichón expulsó un volumen de cenizas comparable al de la erupción del Monte Santa Elena de mayo de 1980, la nube estratosférica que creó era, en promedio, 100 veces más densa. La verdad es que constituye la nube más densa observada en el hemisferio Norte desde la erupción del Krakatoa, en Indonesia, en 1883. La nube arrojada por el Monte Santa Elena se componía principalmente de cenizas finas, que se aglomeraron en seguida para formar partículas mayores que sedimentaron; El Chichón dejó una densa niebla de gotículas de ácido sulfúrico que tardarán varios años en sedimentar completamente.

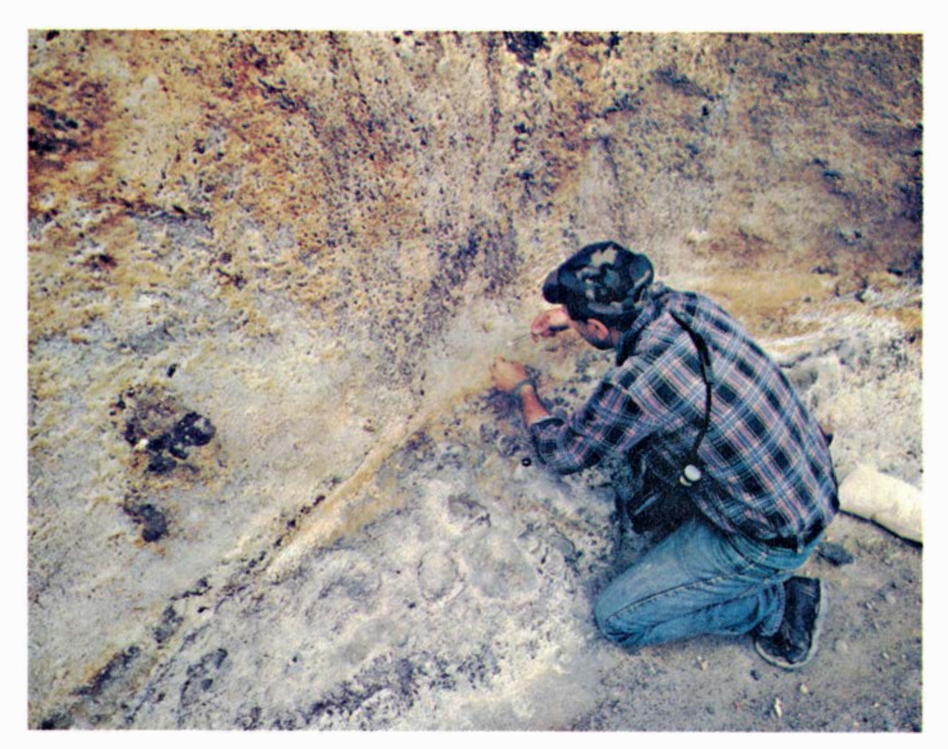
El tamaño de la nube de aerosol puede relacionarse con las características geológicas del volcán, tales como su tipo general de composición y su estilo eruptivo. El Chichón emitió gases sulfurosos suficientes para formar cantidades sustanciales de ácido sulfúrico, y fue lo bastante poderoso como para inyectar los gases en la estratosfera, en parte porque se trataba de un volcán de contenido silíceo intermedio. Los volcanes con bajo contenido silíceo tienden a mostrar un alto contenido de azufre; tienden también a producir erupciones menos explosivas que los dotados de alto contenido de sílice. El Chichón ofrecía, sin embargo, un contenido de azufre mucho más alto que los volcanes del mismo tipo general de composición, anomalía que los geólogos están todavía tratando de explicar. El azufre puede provenir de depósitos sedimentarios subyacentes al volcán o de depósitos de sulfuros sitos en una placa de la corteza terrestre que está zambulléndose debajo de otra, fundiéndose al mismo tiempo.

El aerosol de ácido sulfúrico se forma por reacción fotoquímica entre los gases de azufre, desprendidos por un volcán, y el vapor de agua de la estratosfera. El aerosol causa un descenso de la temperatura media global porque las gotas absorben radiación solar y, además, retrodifunden parte de ella hacia el espacio. Pero las reacciones mediante las cuales se sintetiza ácido sulfúrico son complicadas; no se conocen bien los distintos pasos de las mismas. Tampoco se sabe hasta qué punto los efectos climáticos de una nube de aerosol volcánico dependen de ciertas variables: velocidad de formación del ácido sulfúrico, velocidad de crecimiento de las gotículas de aerosol y velocidad de propagación de la nube. Pobreza de conocimientos que se debe, en parte, a la falta de datos sobre nubes volcánicas de la estratosfera.

Los efectos atmosféricos de El Chichón se han estudiado mejor que los de cualquier otro volcán que generase una gran nube estratosférica. La erupción sirve, pues, de experimento natural, en el que se pueden someter a prueba los instrumentos de seguimiento y vigilancia y los modelos teóricos de la química y de los efectos climáticos de los aerosoles volcánicos en esa zona de la atmósfera. El ingente volumen de información recogido durante y después de la erupción está en fase todavía de análisis. Pero ya se adivinan algunas sorpresas. Por ejemplo, un análisis de los datos de temperatura sugiere que la nube de aerosol pudo haber producido su máximo efecto sobre las temperaturas del hemisferio Norte dos meses des-



1. LA ERUPCION DE EL CHICHON de 1982 reventó los 200 metros superiores del antiguo cono volcánico. El volcán, situado en una remota región del sur de México, había permanecido inactivo unos 600 años. La fotografía del volcán antes de la erupción de 1982 (*arriba*) muestra el borde del antiguo cráter, la cúpula que coronaba la chimenea central y una cúpula lateral (*parte inferior izquierda*) que tapaba una chimenea secundaria. El cráter que dejó la erupción de 1982 (*abajo*) está dentro del cráter antiguo, que queda enmascarado por los nuevos depósitos volcánicos. Durante la estación lluviosa, se acumuló agua en el cráter formando un lago ácido y caliente; la temperatura del agua es de unos 50 grados Celsius; el *p*H, que mide la acidez, es inferior a 1.



2. CONTENIDO DE AZUFRE de los depósitos volcánicos en la linde del cráter; nos lo indica su color amarillo. El análisis de las cenizas volcánicas de la erupción de El Chichón arrojó valores muy altos de sulfato: hasta dos por ciento en peso. El azufre está presente en varias formas; por ejemplo, en cristales de anhidrita ($CaSO_4$), rara en rocas volcánicas. La corteza amarilla en esos depósitos es azufre elemental, que se está probablemente liberando por la descomposición de la anhidrita. Las estimaciones de la masa de los gases sulfurosos inyectados en la atmósfera por la erupción se elevan hasta 20 millones de toneladas.

pués de la erupción, mucho antes que lo predicho por la mayoría de los modelos climáticos.

El Chichón está a 17,33 grados Norte de latitud y a 93,2 de longitud Oeste, en el estado de Chiapas, el más meridional de México. Tres placas de la corteza terrestre convergen en la vecindad del sur de México: norteamericana, caribeña y de Cocos. La actividad volcánica está ligada al hundimiento de la placa de Cocos bajo la placa norteamericana. Sin embargo, El Chichón y otros centros volcánicos asociados están situados en un hueco entre el cinturón volcánico transmexicano al Norte y el cinturón volcánico guatemalteco al Sur. Esta posición aislada quizá tenga que ver con una rotura en la placa de Cocos, causada por el hundimiento de la Dorsal de Tehuantepec en el Pacífico, más allá de la costa sudoeste de México.

El Chichón, descrito por primera vez en 1928, era un pico volcánico que se alzaba a 1260 metros sobre el nivel del mar. Las chimeneas de su interior permanecían cerradas por cúpulas o tapones de lava solidificada. Aunque durante muchos años hubo fumarolas activas en el cráter, el volcán había estado inactivo largo tiempo. Una datación reciente con carbono 14 indica que la última erupción, cuya magnitud decuplicó tal vez la de 1982, se produjo en la segunda mitad del siglo XIV.

La erupción de 1982 vino precedida por un mes de actividad sísmica registrada por una red de sismógrafos instalados en la zona en 1980 para vigilar la posible actividad sísmica causada por el agua embalsada en las presas. Los focos de los terremotos eran inicialmente superficiales (a menos de cinco kilómetros de profundidad); antes de la erupción se hicieron más superficiales todavía (a menos de dos kilómetros). La actividad sísmica debióse, probablemente, al movimiento del magma hacia la superficie o por la interacción de aquél con aguas subterráneas.

A partir de los depósitos formados por el volcán y su correlación con los informes de testigos oculares se ha determinado la secuencia de la erupción. Cartografiaron los depósitos, poco después de la erupción, Haraldur Sigurdsson y Steven, N. Carey, de la Universidad de Rhode Island, y J. M. Espindola, de la Universidad Nacional Autónoma de México. Las divisiones entre las tres erupciones principales del 28 de marzo, 3 de abril y 4 de abril están señaladas por cambios en el tamaño del grano de los depósitos de cenizas. Esas erupciones fueron de tipo pliniano, así llamado en honor de Plinio el Joven, quien describió la erupción del Vesubio en el año 79 de la Era Cristiana. Las erupciones plinianas se caracterizan por grandes columnas de gases, polvo, cenizas y piedra pómez, vigorosamente impulsadas por la convección hasta alturas de decenas de kilómetros. Estas erupciones, que elevaron los gases y el polvo que se convirtieron en la nube estratosférica, vomitaron la mayor parte de los depósitos. La segunda columna de erupción, que se hundió antes de disiparse, produjo las avalanchas superficiales de gas caliente, cenizas y piedra pómez, llamadas corrientes piroclásticas. Estas dejan depósitos de características bien definidas.

La primera erupción pliniana comenzó a las 23 horas 32 minutos del día 28 de marzo. Duró cinco o seis horas. La segunda, iniciada a las 19 horas 35 minutos del 3 de abril, no fue mera repetición de la primera, sino que arrojó gran cantidad de roca volcánica antigua. La presencia de la roca más antigua en la capa de cenizas de esta erupción indica que la chimenea del volcán se estaba erosionando y ensanchando con el paso del magma. El ensanchamiento de la salida y la concomitante disminución de la velocidad de erupción provocaron el hundimiento de la columna.

Las erupciones plinianas se producen cuando la velocidad de erupción es alta y relativamente pequeña la abertura (de diámetro inferior a los 200 metros). La energía cinética que la materia adquiere al pasar por la chimenea impulsa la parte inferior de la columna, mientras que las corrientes de convección creadas en la atmósfera por la materia caliente alimentan la parte superior. Si decrece la velocidad de erupción, la parte alta de la columna puede adquirir mayor densidad que el aire, lo que determinaría su colapso. La materia todavía caliente caerá alrededor del cráter y se propagará hacia fuera, impulsada por la energía cinética adquirida en la caída y convertida en fluido por los gases calientes y el aire. A medida que la materia piroclástica se aleja, se irá separando en dos partes, la corriente, más densa y de movimiento más lento, y las oleadas piroclásticas, más ligeras y rápidas. Las oleadas piroclásticas saldrán en todas direcciones a partir del volcán, precediendo a las corrientes, canalizadas éstas por la topografía.

Las oleadas generadas por El Chichón destruyeron los bosques de sus laderas y colinas adyacentes hasta un radio de ocho o nueve kilómetros con centro en el cráter. La mayoría de los árboles de esta zona quedaron arrancados de cuajo y carbonizados, ennegrecidos o lijados en su lado enfrentado al volcán. Las temperaturas fueron lo suficientemente altas para producir la

combustión de árboles caídos y la carpintería de las casas, pero no bastaron para incendiar árboles vivos. A pesar de su labor destructora, las oleadas dejaron sólo un manto delgado que se ensancha gradualmente hacia el cráter. Las corrientes piroclásticas siguieron a las oleadas, orientándose por los valles estrechos de los ríos de las laderas del volcán. Las corrientes dejaron depósitos radiales de cenizas y piedra pómez bastante gruesos, con bloques de la densa roca volcánica de hasta un metro de diámetro y troncos carbonizados.

La tercera erupción pliniana comenzó a las 5 horas 22 minutos del día 4 de abril. A imagen de la segunda, arrojó ceniza y piedra pómez con grandes cantidades de rocas volcánicas antiguas. Pero esta vez la columna de erupción pareció haberse disipado sin colapsar. En la proximidad del volcán, los depósitos de la tercera erupción están recubiertos por cenizas de grano fino. Capa de ceniza que se produjo, probablemente, cuando la materia piroclástica ardiente de los ríos de lava entró en contacto en los valles con aguas procedentes de las inundaciones estacionales, produciendo explosiones de materia vaporizada.

El volumen de cenizas y piedra pómez arrojado durante esas erupciones de El Chichón puede considerarse moderado: entre 0,5 y 0,6 kilómetros cúbicos. Por mor de comparación, la erupción de 1902 del volcán guatemalteco de Santa María arrojó 10 kilómetros cúbicos de materia, 20 kilómetros cúbicos la erupción de 1883 del Krakatoa y la erupción de 1815 del Tambora, en Indonesia, más de 175 kilómetros.

La ceniza volcánica expulsada por El Chichón pertenecía al tipo de composición llamada andesita; la arrojan las erupciones de volcanes situados en las zonas donde una placa obliga a otra a hundirse bajo ella (el magma de El Chichón, rico en potasio, se encuadra en el tipo de la traquiandesita). La composición química de los magmas varía desde basalto, pobre en silicatos y rico en hierro, hasta dacita y riolita, ricas en sílice y pobres en hierro, pasando por una composición intermedia de andesita. El azufre es más soluble en magmas ricos en hierro, lo que explica mayor concentración en azufre del basalto que la riolita. Mas para generar cantidades importantes de aerosol estratosférico, la explosividad de la erupción ha de ser tal que eleve los gases sulfurosos a gran altura en la atmósfera; ahora bien, los magmas pobres en sílice tienden a producir erupciones menos explosivas que los ricos en sílice. Las erupciones de materia de composición intermedia son apropiadas para inyectar cantidades apreciables de gases sulfurosos en la estratosfera. El archivo histórico sugiere que los efectos atmosféricos de los volcanes y su tipo general de composición tienden a estar correlacionados. En este orden, el Monte Agung, de Balí, que produjo una gran nube estratosférica cuando entró en erupción en 1963, era un volcán de composición andesítica. La erupción del Krakatoa, primera cuyos efectos atmosféricos se estudiaron intensamente, se distinguió por su magnitud y su composición dacítica.

La ceniza depositada por El Chichón se caracteriza por un contenido en azufre anormalmente alto para cualquier clase de erupción. Johan C. Varekamp, de la Universidad de Wesleyan, y James F. Luhr, de la de California en Berkeley, los primeros geólogos en recoger muestras y analizar la composición química de los depósitos de cenizas, encontraron proporciones de sulfato notables: hasta un dos por ciento en peso. Con el microscopio podían verse cristales libres de anhidrita ($CaSO_4$), un mineral de sulfato de calcio escaso en las rocas volcánicas. Cuando se lavó la ceniza, se halló azufre adsorbido en la superficie de las partículas de ceniza; dicho de otra manera, parte de los gases sulfurosos desprendidos en el curso de la erupción los habían eliminado de la nube volcánica las cenizas.

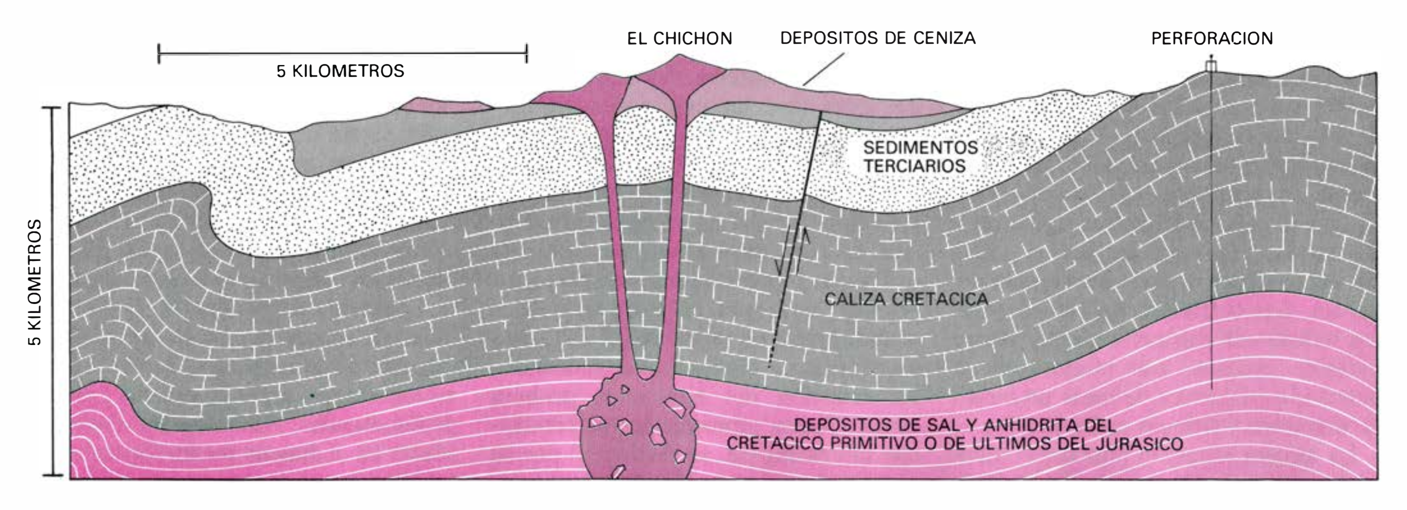
Los niveles anormalmente altos de azufre en las cenizas podrían explicarse con la presencia bajo el volcán de depósitos sedimentarios de minerales ricos en azufre. Una perforación honda, cerca del volcán, realizada en el curso de una exploración en busca de petróleo, había atravesado gruesas capas de anhidrita y sal sedimentarias, formadas por la evaporación de mares poco profundos hace unos 100 millones de años. Para varios investigadores, el magma ascendente hacia el volcán podría haber asimilado grandes cantidades de azufre al atravesar esas capas. El problema del origen del azufre continúa, no obstante, sin resolverse. Según William. I. Rose, Jr., de la Universidad Técnica de Michigan, el azufre habría ascendido con el magma desde grandes profundidades. Notables cantidades de azufre en el magma recién formado parecerían requerir una fuente de azufre muy rica en la placa que se hunde. En algunas de las grietas mesoceánicas, donde se están formando nuevas placas de la corteza terrestre, aberturas hidrotérmicas emiten soluciones calientes ricas en azufre, que recubren la corteza recién creada con depósitos de sulfuros. Cuando tal corteza se sumerge y se funde, se originan magmas de extraordinaria riqueza en azufre.

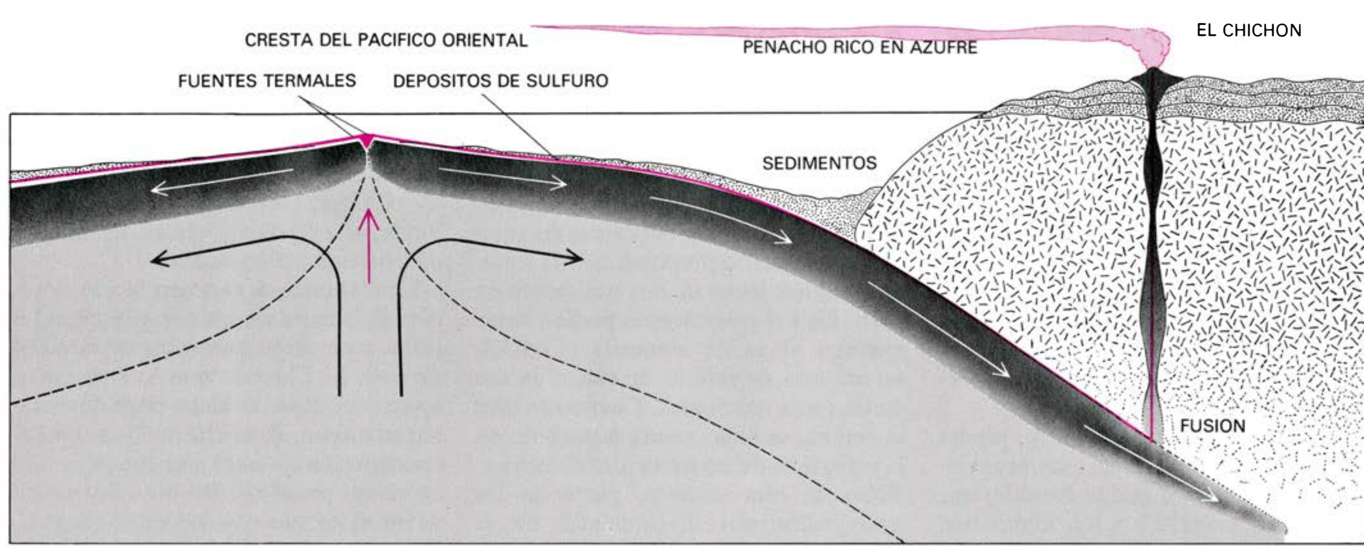
Los estudios de la composición isotópica del azufre en los cristales de anhidrita, y en otros productos de la erupción de El Chichón que lo contienen, aportarán quizá la clave para determinar su origen. El azufre de los sedimentos depositados en el mar será más rico en el más pesado de los dos isótopos de dicho elemento que posee el magma. Las primeras investigaciones hablan de un origen mixto para el azufre expelido en la erupción de El Chichón. No se ha determinado todavía cuál era la contribución exacta de azufre de cada fuente.

En un magma rico en azufre, ¿cuánto de él liberan las rocas fundidas de la erupción? ¿Cómo puede determinar-



3. MAPA TECTONICO del sur de México, que describe la relación de El Chichón con los cinturones volcánicos mexicano y guatemalteco y con características tectónicas locales. Tres placas importantes de la corteza terrestre convergen en esta región. La placa caribeña se desliza por encima de la placa norteamericana a lo largo de un grupo de fallas en Guatemala, mientras se obliga a la placa de Cocos a sumergirse bajo las placas norteamericana y caribeña en la fosa centroamericana, a la altura de la costa de México. El vulcanismo, relacionado con el hundimiento de la placa, presenta ciertos rasgos que nos hablan de una tectónica compleja (compensación entre los cinturones volcánicos trans-mexicano y guatemalteco). La posición aislada de El Chichón puede resultar de una fractura en la placa que se está hundiendo, donde se subduce la Dorsal de Tehuantepec, zona de fracturas con la que está alineada.





4. SE HAN PROPUESTO DOS FUENTES para el azufre lanzado por la erupción de El Chichón. Una profunda perforación cerca del volcán permitió sacar depósitos sedimentarios de anhidrita y sal formados por evaporación de mares someros hace unos 100 millones de años (*arriba*). El magma pudo haber asimilado grandes cantidades de azufre al atravesar esas capas, en su camino hacia la superficie. El corte geológico vertical se funda en el trabajo de Robert I. Tilling y Wendell A. Duffield, del Servicio de Inspección Geológica. Cabe, asimismo, que el azufre estuviera en la roca que se fundió para formar el magma (*abajo*). En algunas porciones de los valles de fractura submarinas, donde se origina nueva corteza oceánica, hay aberturas que desprenden soluciones ricas en azufre. Cuando las soluciones precipitan, recubren la corteza recién formada con depósitos de sulfuro. Una placa portadora de tales depósitos pudo producir un magma rico en azufre cuando se vio forzada a hundirse y fundirse. El azufre del magma arrojado puede provenir de ambas fuentes.

se la fracción del elemento desprendida en forma gaseosa? Joseph D. Devine, de la Universidad de Rhode Island, y Sigurdsson han desarrollado un método para medir la cantidad de azufre en un magma antes de la erupción, a partir de la composición de las inclusiones en cristales. Los cristales que se forman en un magma, inmediatamente antes de la erupción, aprisionan algunas veces, mientras crecen, pequeñas cantidades de la materia fundida en su interior. El magma aprisionado puede formar glóbulos de vidrio, en lugar de cristalizar. La composición del vidrio refleja el contenido preeruptivo en componentes volátiles, tales como los gases ricos en azufre. Devine y Sigurdsson midieron, mediante una microsonda electrónica, el contenido en azufre de las inclusiones fundidas de las cenizas de varios puntos de erupción volcánica. Determinaron así el contenido en muestras de vidrios volcánicos formados en cada erupción. La diferencia entre ambos nos indica el desprendimiento de gases en el curso de la erupción. La cantidad total de gases sulfurosos desprendidos en una erupción puede estimarse a partir de este valor y de medidas del espesor de los depósitos de cenizas y de la superficie que cubren, que dan el volumen total de magma expulsado en la erupción.

Devine y Sigurdsson han hallado una buena correlación entre sus estimaciones de la cantidad de gases sulfurosos desprendidos por los volcanes y la disminución de la temperatura media hemisférica; tal disminución refleja la magnitud de la nube de aerosol que el volcán produjo. Su estimación de la cantidad de azufre desprendido en forma de gas en El Chichón, obtenida a partir de las inclusiones en los cristales, resulta, sin embargo, mucho más baja que los cálculos directos de la cantidad de azufre en la estratosfera después de la erupción. En el caso de El Chichón, parte del azufre en el magma se encontraba en una forma tal que no aparecía en las medidas de gases en las inclusiones de los cristales. El azufre pudo haber quedado aislado antes de la erupción en pequeños cristales de anhidrita, como los hallados en las partículas de ceniza. Durante la erupción, gran parte de este azufre se desprendió quizás en forma de gases sulfurosos.

Los análisis de las emisiones volcánicas muestran que el azufre se emite en forma de dióxido de azufre (SO_2) y, en menor grado, de sulfuro de hidrógeno (H_2S), que no tarda en oxidarse para dar dióxido de azufre. En la

estratosfera, el dióxido de azufre reacciona con los radicales hidroxilo (OH^-) creados por la fotodisociación del vapor de agua, para originar compuestos intermedios –así, el sulfito de hidrógeno (HSO_3)– que, con el tiempo, se condensan dando gotas de ácido sulfúrico (H_2SO_4) y agua. No se sabe exactamente cómo se forman las gotas. El ácido sulfúrico gaseoso se condensa sobre diminutas partículas, tales como el polvo volcánico o incluso iones o grupos de moléculas, que actúan de gérmenes. Las moléculas de sulfito de hidrógeno y agua pueden desarrollar gotículas por simple agrupación; lo mismo cabe decir de las moléculas de vapor de ácido sulfúrico y agua. Las reacciones fotoquímicas son lentas: la conversión completa en aerosol de los gases sulfurosos emitidos dura semanas o meses. A medida que las gotas crecen y precipitan desde la estratosfera, nuevas gotículas continúan creándose. La nube, por consiguiente, se va renovando a sí misma durante algún tiempo.

La dispersión de la nube viene determinada por la configuración de la circulación estratosférica. La estratosfera comprende aquella región entre la baja atmósfera (troposfera) y la alta atmósfera (mesosfera) donde la temperatura del aire no varía mucho con la altura. A causa del débil gradiente de temperatura, hay poca mezcla vertical; los gases volcánicos tienden a quedarse, pues, en la altura hasta donde los inyectó la erupción, formando nubes estratificadas. La estratosfera tiene sus propias variaciones estacionales del tiempo; a finales de primavera y durante el verano del hemisferio Norte, los vientos estratosféricos de la posición subtropical de El Chichón soplan hacia el Oeste. Los vientos estratosféricos transportan rápidamente los aerosoles volcánicos alrededor del globo, pero a éstos les cuesta más extenderse hacia el Norte y hacia el Sur, fenómeno que se ha considerado importante para retrasar el efecto de una erupción volcánica sobre el clima.

Las tres erupciones plinianas de El Chichón inyectaron, por lo que parece, cenizas y gases volcánicos ricos en azufre en la base de la estratosfera subtropical. De las observaciones mediante satélite se desprende que la erupción del 28 de marzo creó una nube a una altura de unos 20 kilómetros, la del 3 de abril otra algo más baja y, la del 4 de abril, una enorme nube centrada a unos 26 kilómetros. Se pudo ver cómo las nubes estratosféricas avanzaban hacia el Oeste, mientras que la nube de cenizas, de corta vida, se extendía hacia el Este en la alta troposfera antes de dispersarse.

La nube estratosférica fue seguida por sistemas telemétricos láser con base en tierra, así como por múltiples instrumentos montados en satélites. Los sistemas telemétricos láser tenían la misión de vigilar los aerosoles atmosféricos; por su parte, algunos de los instrumentos se construyeron para registrar las longitudes de onda a que ciertos componentes gaseosos del aire absorben o emiten radiación. Fue, hasta cierto punto, una sorpresa que, a partir de los datos recogidos, pudiera determinarse la posición de la nube.

El sistema telemétrico láser recibe el nombre de lidar (*l*ight *d*etection *a*nd *r*anging). Mide la cantidad de luz de un pulso láser retrodifundida por la atmósfera; la cantidad de aerosol se deduce comparando la cantidad de luz retrodifundida con la que normalmente lo es por las moléculas de aire de una altura dada. A medida que la nube de aerosol de El Chichón se propagó alrededor del globo, se registraron altos valores de retrodifusión sucesivamente en estaciones lidar con base en tierra, de los Estados Unidos, Japón y Europa.

De cuantas medidas de lidar se realizaron, las más espectaculares correspondieron al Observatorio de Mauna Loa, en Hawai, efectuadas inmediatamente después de la erupción. Hawai está viento abajo de El Chichón; la nube estratosférica no se había dispersado apenas cuando pasó por su vertical. Los valores de retrodifusión registrados por el lidar en ese observatorio fueron los mayores nunca registrados desde que entró en funcionamiento el sistema, en 1973. La nube era 140 veces más densa que la del Monte Santa Elena. En un comienzo, la densidad de la nube permitía su detección con la radiación solar en la región visible del espectro que devolvía al espacio. A medida que se movía alrededor del globo, de Este a Oeste, se la seguía por medio



5. MAPA DE TEMPERATURA de la cima de la nube volcánica el 29 de marzo de 1982, al día siguiente de la primera gran erupción de El Chichón, generado a partir de datos recogidos por el radiómetro avanzado de muy alta resolución, que iba instalado a bordo del satélite *NOAA 7*. Un canal del radiómetro mide la cantidad de radiación que la atmósfera y la Tierra emiten en longitudes de onda del infrarrojo. Antes de la erupción, las temperaturas de la superficie del mar, que pueden servir de base para predicciones a largo plazo del tiempo, se calculaban a partir de los datos del radiómetro. Sin embargo, cuando la nube volcánica pasó bajo el radiómetro el instrumento detectó la emisión de la cima de la nube. El color en esta imagen varía, a medida que sube la temperatura, desde el verde al amarillo y al rojo. El centro de la nube está más frío que la periferia porque se halla más elevado también. Para determinar la altura de la nube, se comparan datos del infrarrojo con perfiles de temperatura obtenidos mediante radiosondas transportadas por globos.

de imágenes generadas a partir de las medidas de reflexión tomadas por muchos satélites, entre ellos los *NOAA 6, NOAA 7, GOES East, GOES West, GMS* (japonés) y *Meteosat* (Europa Occidental). La lengua de aerosol más denso, centrada alrededor de 20 grados de latitud Norte, alcanzó Hawai el 9 de abril, Japón siete días después, el Mar Rojo el 20 de abril, y, el 26 de dicho mes, había dado la vuelta al globo. El movimiento de la nube correspondía a un viento estratosférico del Este, de unos 20 metros por segundo de velocidad, unos 70 kilómetros por hora. En su primera circunvalación, la nube se extendió un poco, hasta cubrir un cinturón de 25 grados, desde unos 5 grados de latitud Norte hasta 30 grados.

Conforme el aerosol se dispersaba, se hizo cada vez más difícil distinguirlo de una nube normal, del vapor de agua, del destello de la luz del sol en las olas oceánicas y de otras fuentes de luz reflejada. Resultó, sin embargo, que algunos instrumentos de los satélites, que no se habían proyectado para detectar aerosoles, pudieron seguirlo durante mucho más tiempo. Por ejemplo, el radiómetro infrarrojo a bordo del *Solar Mesosphere Explorer*, que mide la radiación emitida por el vapor de agua estratosférico a una longitud de onda de 6,3 micrometros, permitió detectar emisiones infrarrojas del aerosol volcánico. Uno de los aspectos más interesantes de los mapas generados a partir de los datos recogidos por este satélite es el de mostrar que el grueso de la nube permaneció al sur de los 30 grados de latitud Norte durante más de seis meses después de la erupción. No se esperaba una situación de bloqueo como ésta en la circulación estratosférica, y resulta tanto más intrigante cuanto que un análisis de los datos de temperatura sugiere que El Chichón produjo su máximo efecto sobre las temperaturas del hemisferio Norte cuando la mayor parte del aerosol estaba todavía confinado a latitudes de un tercio del camino entre el Ecuador y el Polo.

La mayoría de las medidas de la nube hechas por satélites se pueden tomar como meros indicadores de la cantidad de materia en la estratosfera, incluyendo polvo y vapor de agua, además de ácido sulfúrico y aerosol. Sin embargo, poco después de la erupción se descubrió que la cantidad de dióxido de azufre gaseoso en la nube volcánica podía determinarse a partir de medidas tomadas por un espectrómetro para cartografiar el contenido total de ozono, transportado a bordo del satélite *Nimbus 7*. El espectrómetro determina el nivel ozónico de absorción en la región ultravioleta del espectro, comparando la cantidad de luz ultravioleta que la estratosfera está reflejando con un valor obtenido de un modelo teórico de absorción, difusión e intercambio radiativo en la atmósfera.

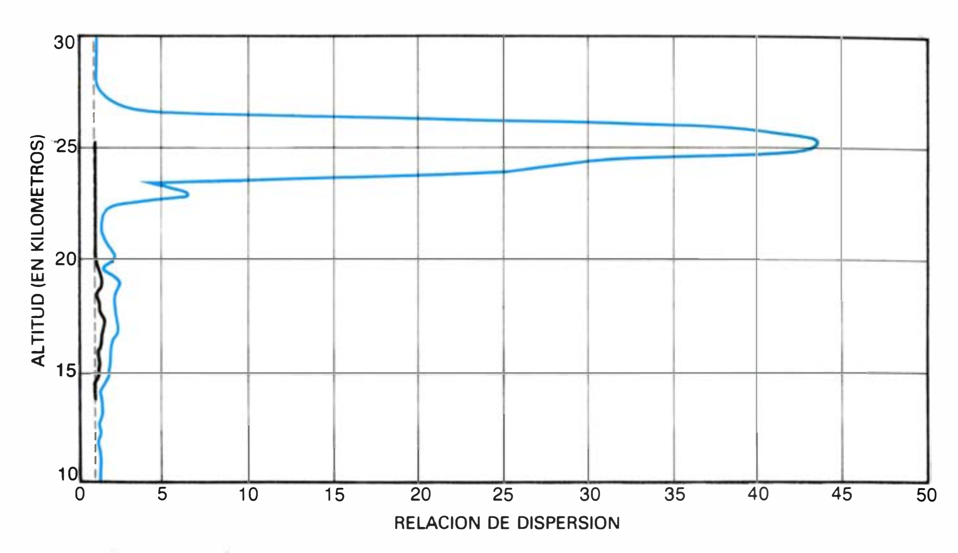
El dióxido de azufre gaseoso reduce la reflectancia de la estratosfera en las longitudes de onda del ultravioleta, lo que indica mayores niveles apa-



6. DIOXIDO DE AZUFRE GASEOSO inyectado en la atmósfera por El Chichón. Se aprecia con nitidez en esta imagen, generada el 5 de abril, un día después de la última de las erupciones mayores, por el espectrómetro para cartografiar la cantidad total de ozono, montado a bordo del satélite *Nimbus 7*. La nube de dióxido de azufre volcánico es la zona negra; se extiende desde Haití hasta el Pacífico, a través de la península de Yucatán. En esta zona de la imagen, el dióxido de azufre de la nube provoca que el espectrómetro registre niveles de ozono falsamente elevados. El espectrómetro mide la cantidad de radiación reflejada por la atmósfera en la región ultravioleta del espectro, donde el principal gas absorbente suele ser el ozono y donde el dióxido de azufre tiene también intensas líneas de absorción. El espectrómetro calcula el nivel de ozono, comparando la cantidad medida con un valor calculado en un modelo atmosférico en el que se introduce información sobre la radiancia atmosférica general. Los colores de la mitad superior de la imagen corresponden a niveles de ozono; varían, con la concentración de ozono, de azul a marrón y a verde. La configuración ondulada en la distribución de ozono (una cresta está centrada sobre los Estados Unidos) es característica de la circulación troposférica superior sobre el norte de los Estados Unidos y Canadá.

rentes de ozono, pero la contribución del dióxido de azufre puede separarse de la del ozono y otros gases si se examina la reflectancia a diferentes longitudes de onda. El dióxido de azufre absorbe en dos longitudes de onda próximas a los 0,3 micrometros, mientras que los otros gases no lo hacen. Basándose en los datos espectrométricos y en estimas de la extensión de la nube, Arlin G. Krueger, del Centro de Vuelos Espaciales Goddard de la NASA, estima que el volcán inyectó en la estratosfera 3,3 millones de toneladas de dióxido de azufre gaseoso. Los datos sugieren también que todo el dióxido de azufre se había convertido en ácido sulfúrico hacia julio de 1982, unos tres meses después de la erupción.

Tal estimación de la cantidad inicial de dióxido de azufre gaseoso es inferior a las realizadas por otros grupos sobre la cantidad de aerosol de ácido sulfúrico en la estratosfera. Esto plantea la pregunta de si el azufre fue inyectado también en forma de otros gases. Una posibilidad es que parte del sulfuro de hidrógeno emitido por el volcán no se oxidó antes de alcanzar la estratosfera.



7. MUESTRAN LOS PERFILES DE DIFUSION de la atmósfera que la erupción de El Chichón produjo una nube estratosférica mucho más densa que la generada en 1980 por la erupción del Monte Santa Elena, que arrojó una cantidad semejante de cenizas, aunque no fue rica en azufre. Los perfiles se construyeron mediante datos obtenidos con el sistema lidar (*l*ight *d*etection *a*nd *r*anging) del Centro de Investigaciones de Langley (adscrito a la NASA), en Virginia, el 18 de julio de 1980 (*en negro*) y el primer día de julio de 1982 (*en color*). La cantidad de materia en la atmósfera se determina por vía indirecta: midiendo la cantidad de luz de un impulso láser que es retrodifundido hasta el instrumento; la altura desde donde ha difundido la luz se establece en razón del tiempo de viaje del impulso difundido. Dicha difusión varía con la densidad de la atmósfera, así como con la cantidad de polvo o aerosol en el aire. La relación aquí representada es el cociente entre la cantidad de luz recibida procedente de una altura dada y la cantidad de luz que las moléculas atmosféricas difunden normalmente. La relación de difusión de 45 (a una altura de 25 kilómetros) es la mayor registrada en el Centro de Investigaciones de Langley.

Las muestras de la nube tomadas por aviones y globos proporcionaron una visión más cabal de su composición y de la dinámica de su crecimiento y dispersión. El aerosol se recogió en vuelos realizados dentro del programa del Centro Espacial Lyndon B. Johnson para recogida de polvo cósmico, así como por un grupo del Centro de Investigación Ames. David S. McKay y Ian D. R. McKinnon, del Centro Johnson, vieron que la nube contenía todavía una apreciable cantidad de ceniza en mayo y principios de junio de 1982. Alrededor del 85 por ciento de la ceniza estaba formada por fragmentos angulares de vidrio. Estos fragmentos estaban recubiertos de gotículas de ácido sulfúrico; algunos de ellos llevaban pequeños cristales que contenían azufre, probablemente en forma de anhidrita u otro mineral de sulfato, sublimado sobre la ceniza durante la erupción o después de la misma.

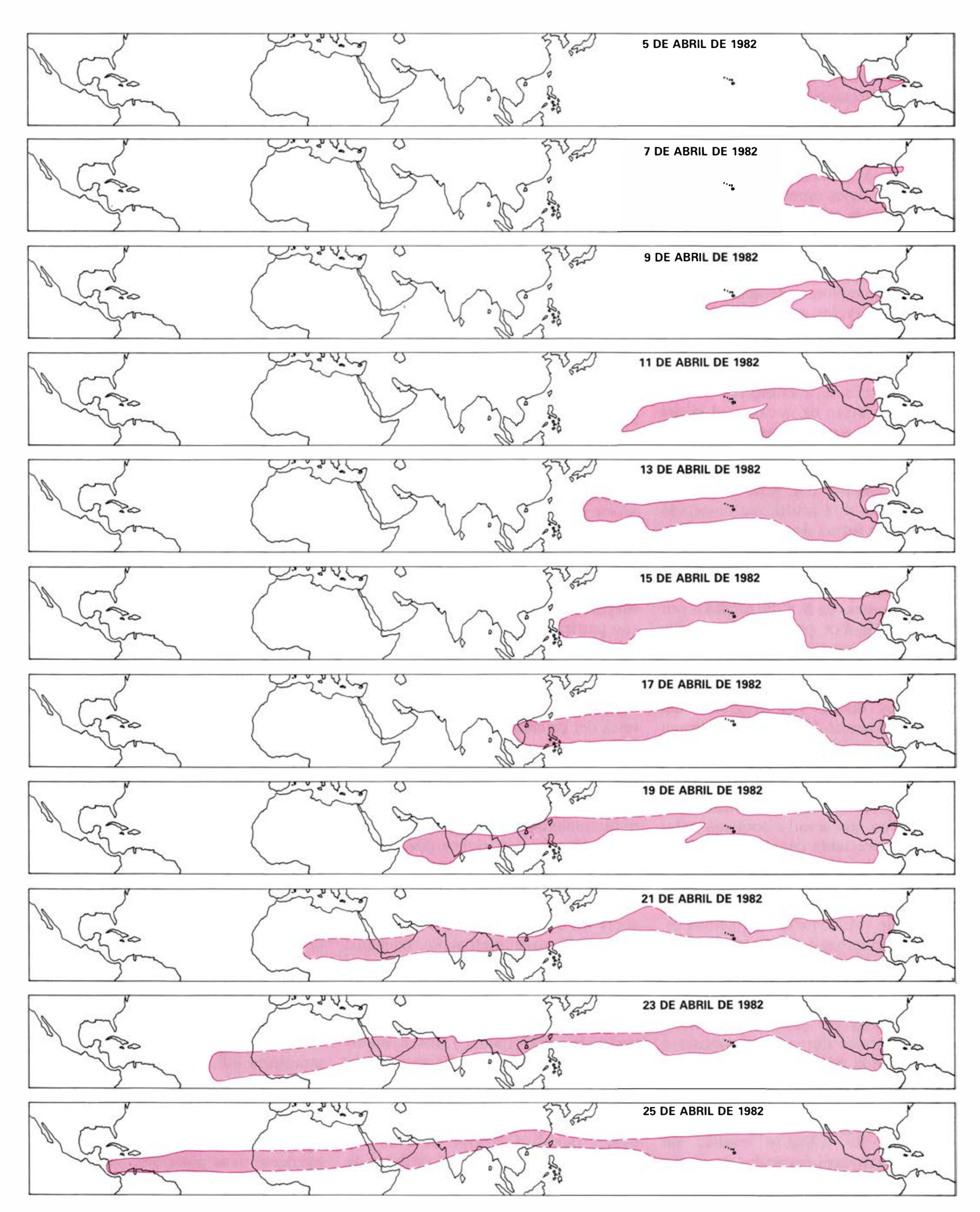
En mayo, las partículas de ceniza presentaban entre tres y seis micrometros de diámetro medio. En julio, las mayores partículas de ceniza se habían depositado fuera de la estratosfera; el diámetro medio se había reducido para entonces a uno o dos micrometros. Las muestras revelaron también que se estaban formando aglomeraciones de ceniza y aerosol de ácido sulfúrico, con densidad suficientemente baja para alcanzar tamaños mucho mayores antes de precipitar. Se hallaron partículas aglomeradas de hasta 80 micrometros de diámetro, con densidades bajas, de 0,1 gramos por centímetro cúbico.

La mejor información sobre la dinámica del crecimiento y eliminación del aerosol en la nube se recogió con instrumentos transportados en globos. David J. Hofmann y James M. Rosen, de la Universidad de Wyoming, han venido tomando muestras de aerosoles estratosféricos con globos desde 1971; elevaron contadores fotoeléctricos de partículas que podían detectar núcleos de condensación (diminutas partículas sobre las que se condensa el vapor) de hasta 0,02 micrometros de diámetro mínimo y distinguir por calentamiento gotas líquidas de partículas sólidas. La representación del tamaño de las partículas de aerosol en la nube en mayo, realizada por Hofmann y Rosen, muestra dos máximos, uno al diámetro aproximado de 0,04 micrometros, debido a gotículas recién formadas, y otro alrededor de 1,4 micrometros, debido a gotículas que han crecido al acumular vapor de ácido sulfúrico. Los datos de posteriores elevaciones de globos en agosto y octubre de 1982 muestran una disminución del número de gotículas pequeñas; al parecer, la producción de núcleos de condensación había cesado hacia esa época. Sin embargo, incluso en octubre había poca disminución del número de grandes gotículas; por lo que se veía, quedaban todavía gotículas en crecimiento reemplazando a las que ya habían precipitado fuera de la estratosfera.

Las medidas realizadas, con globos, de la densidad y tamaño medios de las gotas se han empleado junto con las estimaciones de la extensión de la nube para determinar la cantidad total de aerosol en la estratosfera. De acuerdo con esos datos, un mes después de la erupción podría haber en la estratosfera hasta unos 20 millones de toneladas de aerosol de ácido sulfúrico. En abril de 1983, un año después de la erupción, la cantidad de aerosol había disminuido hasta menos de ocho millones de toneladas.

Los aerosoles estratosféricos afectan el balance global de radiación, porque absorben y retrodifunden radiación solar incidente, aunque también absorben radiación infrarroja emitida por la atmósfera inferior. La absorción y retrodifusión dan lugar a efectos ópticos atmosféricos y a una reducción de las temperaturas en superficie. La absorción del infrarrojo en la radiación solar incidente tiene por efecto principal aumentar las temperaturas estratosféricas.

Después de la erupción de El Chichón, la temperatura de la estratosfera ecuatorial subió unos cuatro grados Celsius; se alcanzó la cota más alta jamás registrada desde que se iniciaron en 1958 las medidas permanentes de la



8. MAPAS GLOBALES de la posición de la nube volcánica durante las tres semanas inmediatas a la erupción de El Chichón. Se construyeron combinando la información de los satélites geoestacionarios *GOES East* y *GOES West* y el de órbita polar *NOAA 7*. La extensión de la nube se determinó examinando las imágenes que los satélites producían a partir de medidas de la radiación visible reflejada. Justo antes de la erupción volcánica, el viento de la estratosfera había empezado a soplar del Este, al cambiar la circulación estratosférica a su configuración de verano. El tiempo que tardó la nube en dar la vuelta al globo terrestre corresponde a una velocidad del viento de unos 20 metros por segundo (unos 70 kilómetros por hora). La nube quedó confinada, sin embargo, a un estrecho margen de latitudes por causa de la configuración de la circulación. El borde de la nube es una línea de trazos donde no se ha podido determinar exactamente. Las nubes volcánicas generadas por otras erupciones en los últimos dos años (entre ellas la del Monte Santa Elena, la de Alaid en la Unión Soviética y la de Galunggung en Indonesia) se dispersaron, en sólo unos pocos días, y no pudieron detectarse en imágenes de satélites.

temperatura estratosférica. Poco después de la erupción empezaron a llegar informes de fenómenos ópticos atmosféricos semejantes a los que siguieron a la erupción del Krakatoa en 1883 y a la del Monte Agung en 1963; se observaban todavía algunos a finales de 1983. Aden B. Meinel y Marjorie P. Meinel, de la Universidad de Arizona, informaron sobre amaneceres y puestas de sol de un fulgor destacado y prolongada duración, que, en Tucson, comenzaron a finales de abril. Se hicieron visibles largas bandas de aerosol en la alta atmósfera el 7 de mayo, cuando el cuerpo principal de la nube atravesaba el noroeste de la ciudad. Durante algún tiempo, el cielo de Arizona adquirió un azul blanquecino, en sustitución del azul brillante que caracteriza al Suroeste de la Unión.

Las puestas de sol que indican la presencia de un aerosol estratosférico comienzan con una luminiscencia verdosa que aparece a gran altura sobre el horizonte y cambia gradualmente a amarillo y anaranjado. Después del ocaso, se produce frecuentemente un crepúsculo rojo fuerte debido a la luz solar reflejada por la nube de aerosol que, más arriba sobre el horizonte, toma un tinte púrpura. La estratificación de la nube se manifiesta a través de rayas horizontales. Varios observadores informaron también haber visto anillos de Bishop, otro efecto óptico típico de los aerosoles volcánicos. Tales anillos, llamados así en honor del reverendo Sereno E. Bishop a quien debemos su descripción tras la erupción del Krakatoa, son halos alrededor del Sol cuyo orden normal de colores está invertido, con el rojo en la periferia.

El efecto más importante que la nube de aerosol de El Chichón producirá sobre el clima tendrá que ver con la reducción de la temperatura media global. Los aerosoles volcánicos se caracterizan por causar una disminución inferior a un grado Celsius. Aunque una caída aparentemente tan pequeña puede resultar significativa para el clima, no será fácil separarla de las fluctuaciones naturales. Ello adquiere particular interés en el caso de la erupción de El Chichón: piénsese que 1981 fue el año más cálido del que se tenga noticia en el hemisferio Norte (aproximadamente 0,5 grados por encima del valor promedio de muchos años) y las temperaturas habían empezado a bajar a principios de 1982, aún antes de la erupción.

Los investigadores que han examinado la relación estadística entre erupciones volcánicas y disminuciones de la temperatura media llegaron al descubrimiento de que las temperaturas promediadas alrededor de un hemisferio descendían unos 0,3 grados Celsius durante un intervalo de uno a tres años después de las erupciones más importantes. Se realizaron predicciones específicas acerca del efecto que la erupción de El Chichón produciría sobre las temperaturas en vista de la variación resultante en el espesor óptico (opacidad) de la estratosfera y a partir de diferentes modelos teóricos del clima. Para calcular la opacidad se compara la cantidad de luz recibida por un fotómetro o un pirheliómetro con los valores normales. Basándose en el descenso de temperatura provocado por anteriores erupciones que dieron lugar a variaciones similares de la opacidad estratosférica, se predijo que El Chichón haría que las temperaturas del hemisferio Norte bajaran alrededor de 0,4 grados hacia fines de 1983. De acuerdo con el modelo climático desarrollado por Alan Robock, de la Universidad de Maryland, se espera un enfriamiento de 0,5 grados para el próximo invierno (el invierno de 1984-1985). Según otros modelos, asistiremos a una disminución similar, aunque el retraso en alcanzarla varía de un modelo a otro por mor de las hipótesis acerca de la expansión y disipación de la nube volcánica.

Un análisis reciente de los descensos de temperatura subsiguientes a erupciones volcánicas (entre ellas la de El Chichón) llegaba a algunas conclusiones sorprendentes. P. M. Kelly y Chris B. Sear, de la Universidad de Anglia Oriental, centraron su trabajo en las temperaturas medias mensuales del hemisferio Norte que siguieron a las principales erupciones de los últimos 100 años. El máximo descenso de temperatura, concluyen, se produce en época temprana, a los dos meses de la erupción en el hemisferio Norte; sucede luego una segunda disminución un año más tarde, después de la cual la temperatura tiende a restablecerse lentamente. La respuesta de las temperaturas del hemisferio Norte a erupciones en el hemisferio Sur puede retrasarse en siete u ocho meses.

El análisis de Kelly y Sear sugiere que la erupción de El Chichón, a 17 grados de latitud Norte, produjo su máximo efecto sobre las temperaturas medias mensuales (un descenso de 0,2 grados Celsius) ya en junio de 1982, a los dos meses de la erupción. Y, como este máximo se presentó antes de que la nube ciñera el globo y portaba todavía cantidades apreciables de polvo, se deduce que una nube de aerosol volcánico afectará la temperatura en un cinturón de latitud mayor que el que la nube cubre, y que tanto el polvo como el ácido sulfúrico contribuyen al enfriamiento.

El descenso de la temperatura en superficie quizá no sea el único efecto de la erupción de El Chichón sobre el clima global. Al alterar el perfil vertical de temperaturas atmosféricas, puede haber variado la configuración de la circulación atmosférica, contribuyendo tal vez al intenso episodio de El Niño del invierno de 1982-83. Por episodio de El Niño se entiende cierto cambio brusco que sufre la circulación atmosférica y la oceánica en el Pacífico ecuatorial. Aunque El Niño extiende su impronta por todo el Pacífico, se le asocia con el cambio súbito de las corrientes oceánicas mar adentro de la costa occidental de América del Sur, su punto de arranque. Allí se debilitan periódicamente los vientos alisios del Sureste. A medida que éstos cesan, el agua del Pacífico occidental aflora hacia el Este, interrumpiendo la Corriente de Humboldt, que fluye hacia el Norte, a lo largo de la costa de América del Sur; esta desorganización produce a su vez un tiempo anómalo sobre el continente.

La nube de aerosoles pudo provocar ese episodio de El Niño al calentar la atmósfera superior en los trópicos. El calentamiento de la estratosfera reduciría la diferencia de temperatura entre ella y la superficie de la Tierra, lo que podría haber originado un debilitamiento de la circulación atmosférica y de la circulación oceánica que aquella impulsa. El episodio de 1982-83 llamó la atención al no seguir el curso normal de desarrollo. Tuvo un comienzo anómalo (se inició en mayo, sin esperar a los habituales octubre o noviembre) y modificó el aspecto de la circulación en la totalidad del Pacífico al mismo tiempo aproximadamente, en lugar de moverse desde la costa de América del Sur hacia el Oeste.

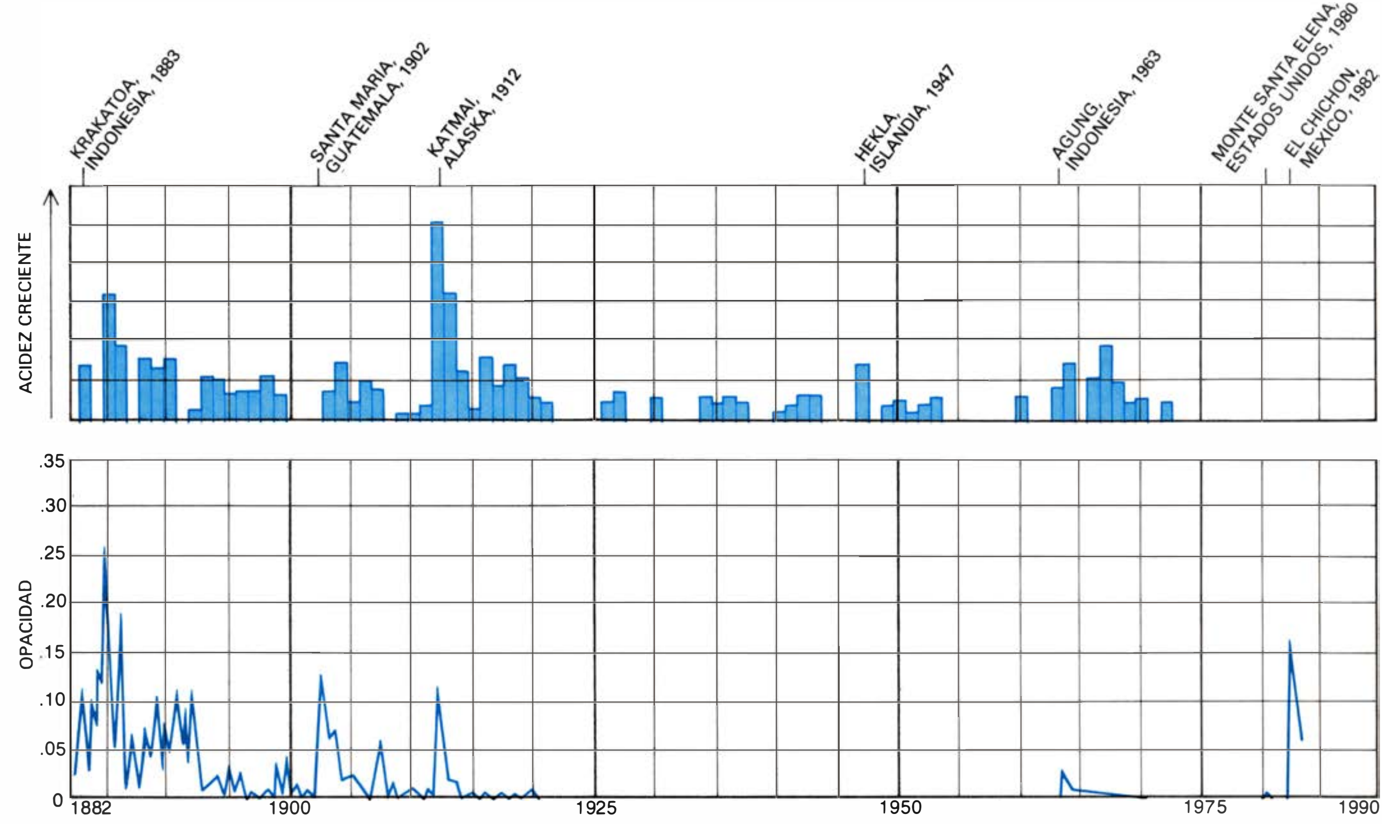
De los nueve episodios que ha contabilizado dicha corriente desde 1950, sólo dos han escapado del curso normal: el de 1982-83 y el de 1963, que coincidió con la erupción del Monte Agung, la última erupción que lanzó una densa nube de aerosol que afectó a las latitudes tropicales. Por otro lado, determinados fenómenos que preceden a un episodio de El Niño, tales como la caída de presión atmosférica en Tahití y en la Isla de Pascua frente a las presiones de Australia e Indonesia, parecen

haber comenzado antes de la erupción de El Chichón. Por ahora, no hay pruebas de una relación causal entre erupciones volcánicas y episodios atípicos de El Niño; probablemente tendremos que esperar futuros episodios para resolver el problema.

Ahora que El Chichón ha establecido la importancia del contenido de azufre de una erupción volcánica en relación con su impacto atmosférico, ¿será posible volver atrás y establecer la relación entre variaciones climáticas y contenido de azufre de erupciones volcánicas pasadas? Desgraciadamente, las erupciones andesíticas como la de El Chichón dejan poco rastro geológico. Los volcanes andesíticos tienden a producir erupciones con escaso magma y dejan unos delgados depósitos de ceniza que se erosionan con facilidad. Será por tanto difícil, si no imposible, que los geólogos reúnan un archivo completo de erupciones ricas en azufre a partir del estudio de depósitos de ceniza. Ahora bien, las erupciones pequeñas y frecuentes pueden producir un efecto sobre el clima tan importante como el de una gran erupción. La investigación de modelos climáticos sugiere que los períodos de actividad volcánica exacerbada podrían explicar, en principio, algunas de las disminuciones globales de temperatura a plazo bastante largo.

Un método prometedor para detectar antiguas erupciones ricas en azufre consiste en medir la acidez de las capas anuales de hielo en los grandes bancos terrestres de hielo (de Groenlandia y la Antártida). La cantidad total de aerosol de ácido sulfúrico en la atmósfera puede calcularse a partir de la acidez de una capa de hielo, si se conoce la latitud de la erupción; adviértase, empero, que los volcanes próximos al banco de hielo pueden producir un efecto desproporcionadamente grande en la acidez porque sus gases sulfurosos llegan a través de la troposfera. La comparación de los datos climáticos con la acidez de las capas de hielo en glaciares ha puesto al descubierto interesantes correlaciones. Así, durante el período entre aproximadamente 1350 a 1700 de la Era Cristiana, conocido por Pequeña Era Glaciar, los hielos que se depositaron eran bastante ácidos. Las correlaciones sugieren que el vulcanismo rico en azufre puede constituir un importante modulador de las tendencias a largo plazo del clima global.

Se está todavía analizando la ceniza depositada por El Chichón, así como tomando muestras del aerosol estratosférico mediante globos, aviones y satélites. (Los instrumentos instalados en tierra siguen detectándolo.) La información ha ayudado ya a profundizar en el conocimiento de la relación entre volcanes y clima global; los ulteriores análisis de los datos y sus aplicaciones al modelado y predicción del efecto de las erupciones volcánicas sobre el clima ocuparán, los próximos años, a los científicos atmosféricos y a los vulcanólogos. La erupción proporcionó un ejemplo de prueba que hacía mucha falta para comprobar la hipótesis de que el contenido de azufre de una erupción produce, sobre el tamaño de la nube de aerosol que crea, un efecto más importante que el volumen de ceniza expulsado. Aunque quedan por ver aquí a qué conclusiones conducirán los futuros análisis, la erupción de El Chichón ha suministrado un aluvión de datos sobre los efectos de esos episodios en el clima mundial sin parangón hasta ahora.



9. MAXIMOS DE ACIDEZ en las capas de hielo del casquete de Groenlandia. Se corresponden con períodos en los que se observó una atmósfera menos transparente que de costumbre. La acidez de las capas anuales de una columna de hielo tomada del casquete a principios de los años setenta (*arriba*) constituye una buena referencia de la cantidad del aerosol de ácido sulfúrico en la estratosfera en cada época. El archivo sobre la opacidad de la estratosfera en el hemisferio Norte (*abajo*) se compiló con datos tomados sobre la intensidad de la luz solar o de las estrellas en los observatorios. La mayor parte de los máximos se puede emparejar con erupciones contemporáneas, de las que se indican algunas (*en la parte superior de la figura*). El período transcurrido entre 1920 y 1960 conoció, en proporción, escasas macroerupciones explosivas, que son las que producen nubes estratosféricas. Al interpretar los archivos hay que atender a la latitud de una erupción volcánica. La del Katmai, en 1912, y la del Hekla, que se produjo en 1947, pudieron producir un gran efecto sobre la acidez del hielo groenlandés a causa de su alta latitud. La acidez de la columna de hielo fue medida por Claus U. Hammer, de la Universidad de Copenhague. El archivo de la opacidad hasta 1970 lo compilaron Owen B. Toon y James B. Pollack, adscritos al Centro de Investigación Ames.

Colisiones a alta energía entre núcleos atómicos

Se provoca el choque entre núcleos pesados, que viajan al 95 por ciento de la velocidad de la luz, y núcleos en reposo. Entre los residuos hay anomalones, fragmentos misteriosos que vuelven a colisionar, mucho antes de lo esperado

Wm. C. McHarris y John O. Rasmussen

En el primer vuelo tripulado a la Luna, en 1969, los astronautas que iban a bordo del *Apolo 11* informaron sobre un fenómeno curioso: cuando cerraban los ojos para dormir, veían a veces un destello puntual de luz. La interpretación que no tardó en dársele en la Tierra lo asignaba a núcleos de átomos pesados, existentes en los rayos cósmicos que incidían contra la nave. La energía depositada por cada núcleo sobre la retina del astronauta era aparentemente mayor que el mínimo necesario para estimular las células de la retina sensibles a la luz.

Poco después del vuelo se comprobó la hipótesis. En el Lawrence Berkeley Laboratory de la Universidad de California, el Bevatrón, un acelerador de protones de alta energía construido a comienzos de la década de los cincuenta, había sufrido una reciente modificación al objeto de que acelerara núcleos más pesados que los solos protones. Edwin M. McMillan, director del laboratorio por aquel entonces, puso voluntariamente su cabeza en la trayectoria de un haz de baja intensidad de núcleos pesados de alta energía. Vio la misma clase de destellos de que hablaban los astronautas.

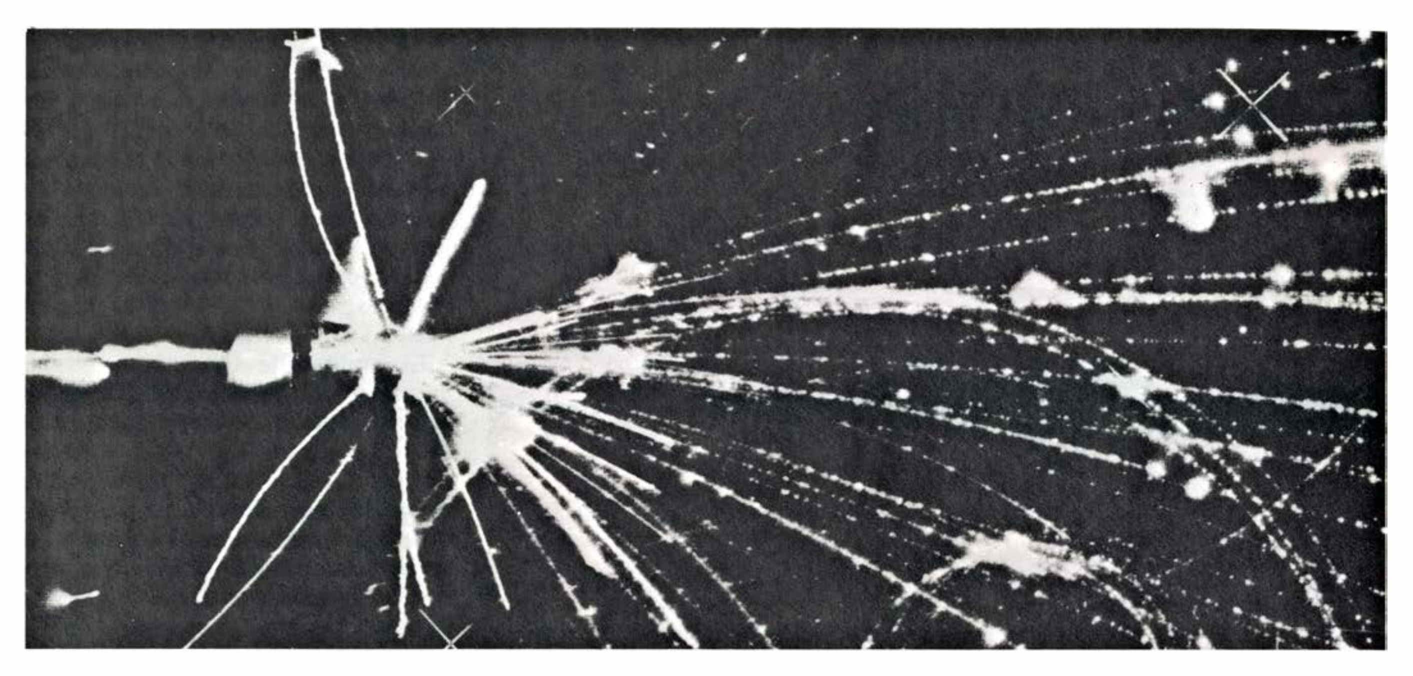
De entonces acá, las colisiones núcleo-núcleo de alta energía han ido mostrando propiedades más difíciles de explicar. Originan, por ejemplo, los llamados anomalones: fragmentos nucleares que colisionan de nuevo, tras una distancia que sorprende por su cortedad. Se puede decir que algunas propiedades ponen de manifiesto una "fuga" de la fuerza que enlaza los nucleones (protones y neutrones) en un núcleo. Las colisiones pueden revelar la actividad de los quarks, constituyentes teóricos de protones y neutrones, no vistos hasta ahora. Parece cierto que la temperatura y presión extremas generadas por las colisiones núcleo-núcleo de alta energía crearán nuevos estados de materia nuclear, como el estado del universo en sus primeras millonésimas de segundo de existencia.

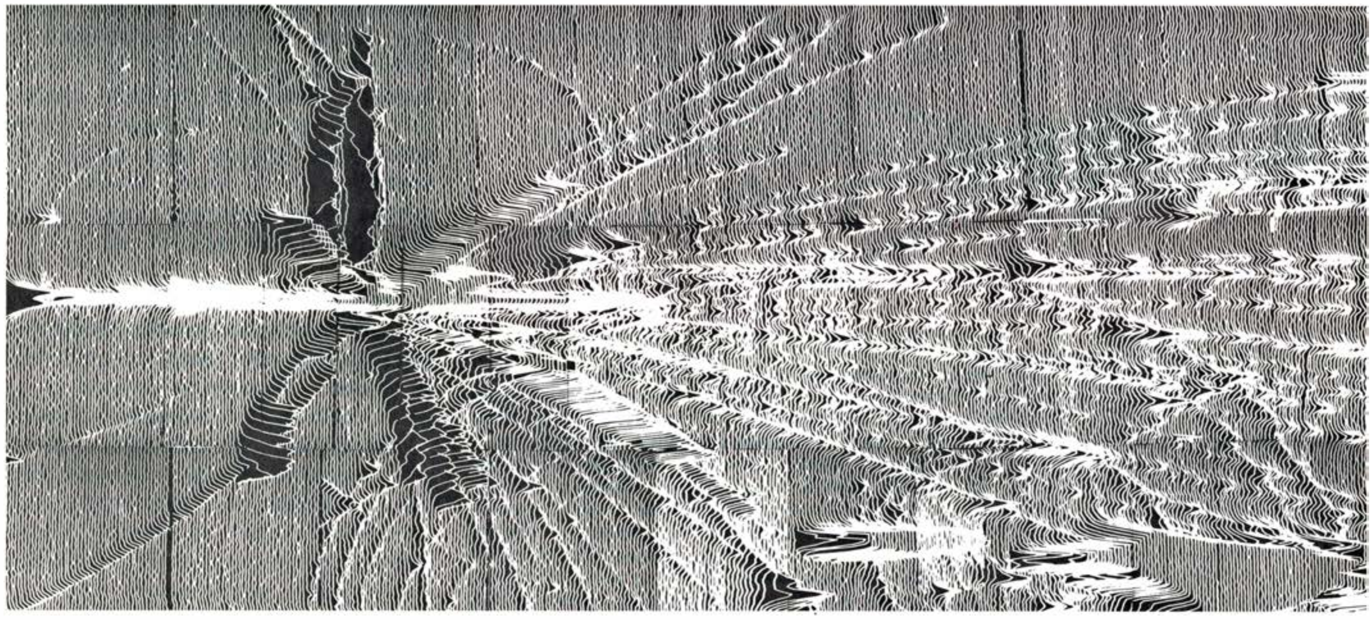
¿Qué caracteriza a las colisiones núcleo-núcleo de alta energía? ¿Por qué los destellos que causan en el ojo se descubrieron en el espacio y no en cualquier pasajero de avión con ganas de dormitar? La diferencia más importante reside en la cadencia o ritmo con que la colisión deposita energía en la materia. La cadencia es proporcional al cuadrado de la carga eléctrica del núcleo "proyectil" que participa en la colisión. Así, un núcleo de hierro de los rayos cósmicos –que, despojado de los electrones con los que formaría un átomo de hierro, se convierte en un ion con carga positiva de 26 unidades– deposita su energía a una cadencia de 26^2, es decir 676, veces mayor que lo haría un protón de los rayos cósmicos de igual velocidad. Además, los núcleos pesados de los rayos cósmicos tienden a romperse cuando colisionan con los núcleos de los gases de la región superior de la atmósfera. Ello asegura que los núcleos pesados de los rayos cósmicos no penetren en las zonas bajas de la atmósfera.

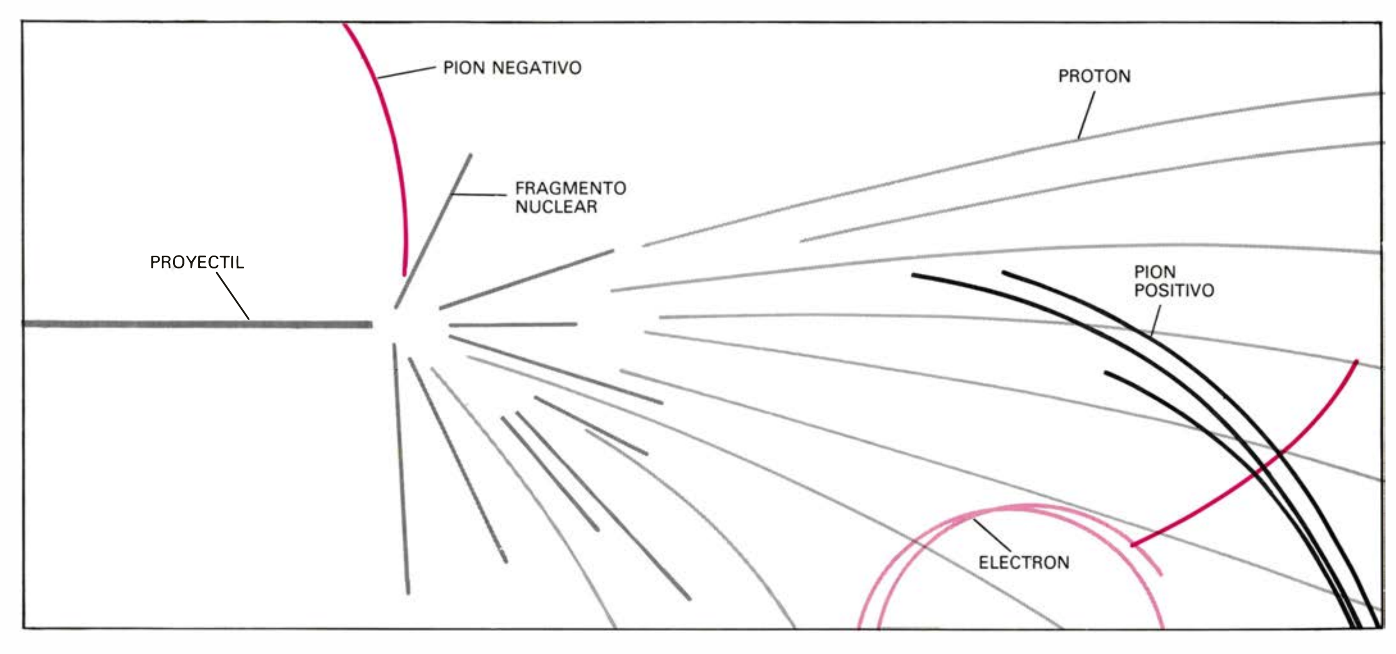
Mucho antes del vuelo del *Apolo 11*, la cadencia con que los iones pesados depositan energía había conducido a los investigadores médicos a recomendar los aceleradores de iones pesados de alta energía. Razonaban que bastaba un impacto de tal ion para matar una célula cancerosa; iba más allá de la mera interrupción de su actividad. (El daño biológico ocasionado por la mayor parte de las radiaciones se debe a la producción de peróxidos y otros radicales químicos en las células.) A principios de la década de los setenta, había en funcionamiento varios aceleradores de iones pesados, pero sus haces tenían un alcance de escasos milímetros en los tejidos. Acelerar los iones a mayores energías exigiría recurrir a los anillos en los que trabajan los físicos.

Eran los tiempos en que los físicos se mostraban reacios a cerrar viejos aceleradores, el Bevatrón entre ellos. Pero se les vio un nuevo servicio: se construyó una línea de transferencia de vacío de manera que el acelerador lineal de iones pesados (HILAC), también en Berkeley, mandara sus haces al Bevatrón. Los iones que salían del HILAC con una energía de 8,5 MeV por nucleón (es decir, 8,5 millones de electronvolt por cada protón y neutrón del núcleo) podían ser de nuevo acelerados a energías de hasta 2 GeV (2000 millones de electronvolt) por nucleón. La

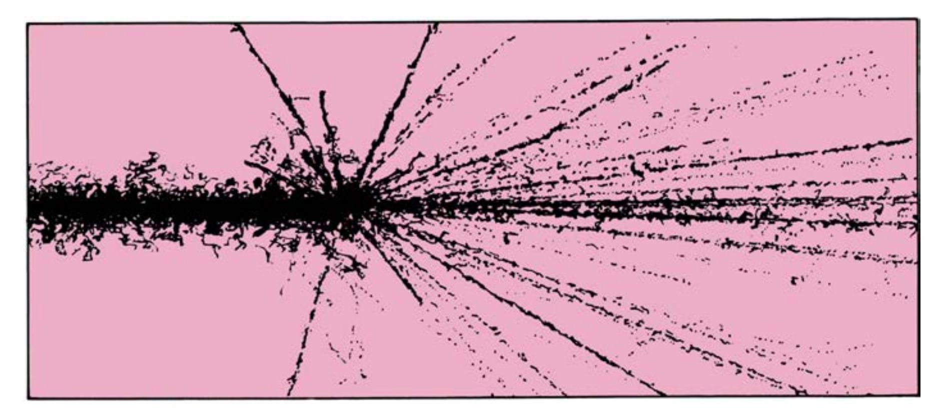
1. COLISION NUCLEO-NUCLEO (*arriba*), fotografiada en una cámara de "cintas" en el Lawrence Berkeley Laboratory de la Universidad de California. En la cámara, el paso de una partícula cargada deja tras de sí una traza de gas ionizado. Un instrumento acoplado a la carga (una cámara de televisión controlada por computadora) reconstruye la colisión (*centro*). Los picos indican la intensidad de la luz, que representa, a la vez, la densidad de ionización de las trazas y la distancia de cada traza a la cámara. Un esquema (*abajo*) identifica algunas partículas. El núcleo de un átomo de argón 40 entró por la izquierda. Su energía cinética era de 72.000 millones de electronvolt (72 GeV). A unos 0,3 metros hacia el interior de la cámara colisionó contra el núcleo de un átomo de plomo de un blanco de óxido de plomo. La mayoría de los residuos detectables despedidos hacia adelante en la colisión eran protones; sus trayectorias se curvaron ligeramente hacia abajo, por efecto de un campo magnético de la cámara. Había también algunos piones residuales, creados por la energía de la colisión. De entre estos últimos, algunos positivos cruzaron las trayectorias de los protones; un pion negativo retrocedió hacia la izquierda. Los fragmentos más pesados eran agregados de protones y neutrones. Dejaron trazas cortas y altamente ionizadas.

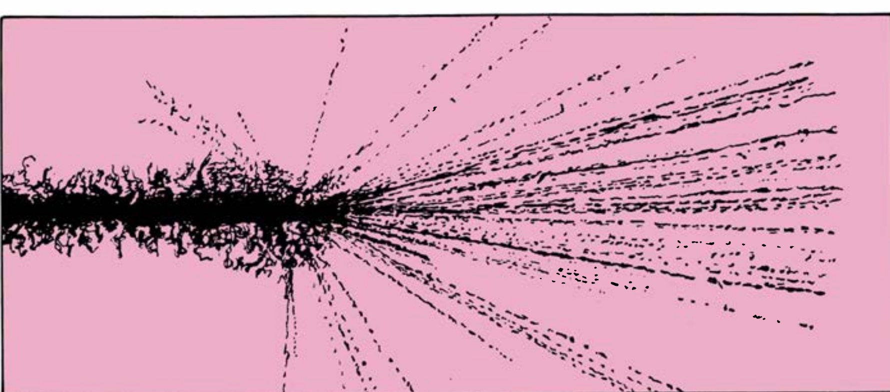





PION NEGATIVO
PROTON
FRAGMENTO NUCLEAR
PROYECTIL
PION POSITIVO
ELECTRON







2. TRES TIPOS DE COLISIONES núcleo-núcleo de alta energía. Se registraron, también en el Lawrence Berkeley Laboratory, colocando emulsiones fotográficas en la trayectoria de un haz nuclear de uranio 238, después de que se acelerase, cada núcleo, hasta una energía de 228 GeV. En una colisión periférica (*arriba*), un núcleo de uranio pasó tan cerca de un núcleo de la emulsión que ambos interaccionaron, si bien sólo electromagnéticamente. La interacción hizo que el uranio se fisionara en dos fragmentos nucleares. El núcleo blanco permaneció intacto (los rizos que salen de las trayectorias se deben a electrones energéticos arrancados a numerosos átomos de la emulsión). En una colisión híbrida (*centro*), un núcleo de uranio choca contra un núcleo de la emulsión, aunque no frontalmente; ambos se desintegran. Algunas de las trazas más oscuras se deben a fragmentos del núcleo blanco, despedidos a grandes ángulos, incluso hacia atrás. En una colisión central (*abajo*), un núcleo de uranio alcanzó de lleno un núcleo de la emulsión. Después, el aglomerado se desintegró y su momento lanzó la mayor parte de los fragmentos hacia adelante. Las trazas correspondientes a cada colisión se han aumentado aquí unas 400 veces.

combinación de ambas máquinas, conocida por Bevalac, podía por vez primera acelerar núcleos pesados, en el laboratorio, hasta velocidades altísimas, del 95 por ciento de la velocidad de la luz; y, con ello, conferirles energías próximas a las que caracterizan a los rayos cósmicos. Se dice que estas velocidades y energías son "relativistas" porque los cuerpos que se mueven a una fracción apreciable de la velocidad de la luz se comportan según lo predicho por la teoría de la relatividad restringida. Mientras, en el Instituto de Investigación Nuclear de Dubna, en la URSS, otra máquina, el Sincrofasotrón, empezó a acelerar núcleos de carbono al doble de la energía que podía darles el Bevalac. El haz era, no obstante, menos intenso.

Simultáneamente a estos avances técnicos, la física teórica daba un gran salto hacia el mejor conocimiento de la composición de la materia: la teoría de los quarks. Ese modelo de las partículas nacía con la pretensión de explicar la estructura interna de protones y neutrones. Se sabía ya que los núcleos atómicos constaban de protones y neutrones situados en un conjunto de "capas" orbitales. Ahora se afirmaba que protones y neutrones estaban formados, cada uno, por tres quarks. Constaba el protón de dos quarks "arriba" (*up*) y un quark "abajo" (*down*); el neutrón, de dos quarks "abajo" (*down*) y un quark "arriba" (*up*). Los quarks tienen carga eléctrica fraccionaria; en unidades de la carga del protón, la carga del quark abajo es $-1/3$ y, la carga del quark arriba, $+2/3$. La teoría de los quarks había ya explicado algunas cuestiones fundamentales relativas a la mezcla de estados que aparece en la combinación de un protón y un neutrón para formar un deuterón (núcleo de hidrógeno-2). Además, como veremos, la teoría de los quarks había puesto en claro algunas características, inexplicables hasta entonces, de las colisiones núcleo-núcleo a alta energía.

Es gratificante observar cómo las nuevas posibilidades de producción de haces de núcleos pesados de alta energía llenan un vacío en nuestro conocimiento. Sea un gráfico que represente la energía que puede darse a las partículas, en el laboratorio, en función de su masa. Los físicos de altas energías han progresado a lo largo del eje de energía, empleando haces de partículas individuales: generalmente haces de electrones, protones y, también, mesones, la clase de partículas de masa intermedia entre la del electrón y la del protón. Por su parte, los físicos nucleares han progresado a lo largo del eje de masa. Esto deja la zona distante de ambos ejes cual "terra incognita", en espera de ser explorada. Examinemos lo que ya se sabe que está ahí.

Las colisiones núcleo-núcleo a energías relativistas pueden clasificarse en tres categorías: periféricas, híbridas y centrales. Aunque la división es un tanto arbitraria, cualquier colisión puede asignarse, en principio, a una de esas categorías. En las colisiones periféricas, un núcleo, el "proyectil", roza un segundo núcleo, el "blanco", o pasa suficientemente cerca de él; así, ambos interaccionan, por lo menos, eléctricamente. Gran parte del proyectil continúa hacia adelante. En las colisiones híbridas, el proyectil golpea al blanco; no lo hace de frente, sino arrancando la mitad del blanco, por ejemplo. Se desintegran fragmentos extensos de ambos núcleos. En las colisiones centrales, el proyectil golpea el blanco más o menos frontalmente. La amalgama de la conjunción de ambos se desintegra mientras que el momento del proyectil tiende a dirigir los fragmentos residuales hacia adelante.

Las colisiones periféricas merecen un examen detallado. El resultado de las colisiones se registra en detectores de diverso tipo: emulsiones fotográficas, donde el paso de una partícula cargada a través de una hoja de bromuro de plata deja una traza de plata metálica; cámaras de burbujas, donde el paso de una partícula cargada a través de un líquido, sobrecalentado bajo presión, da

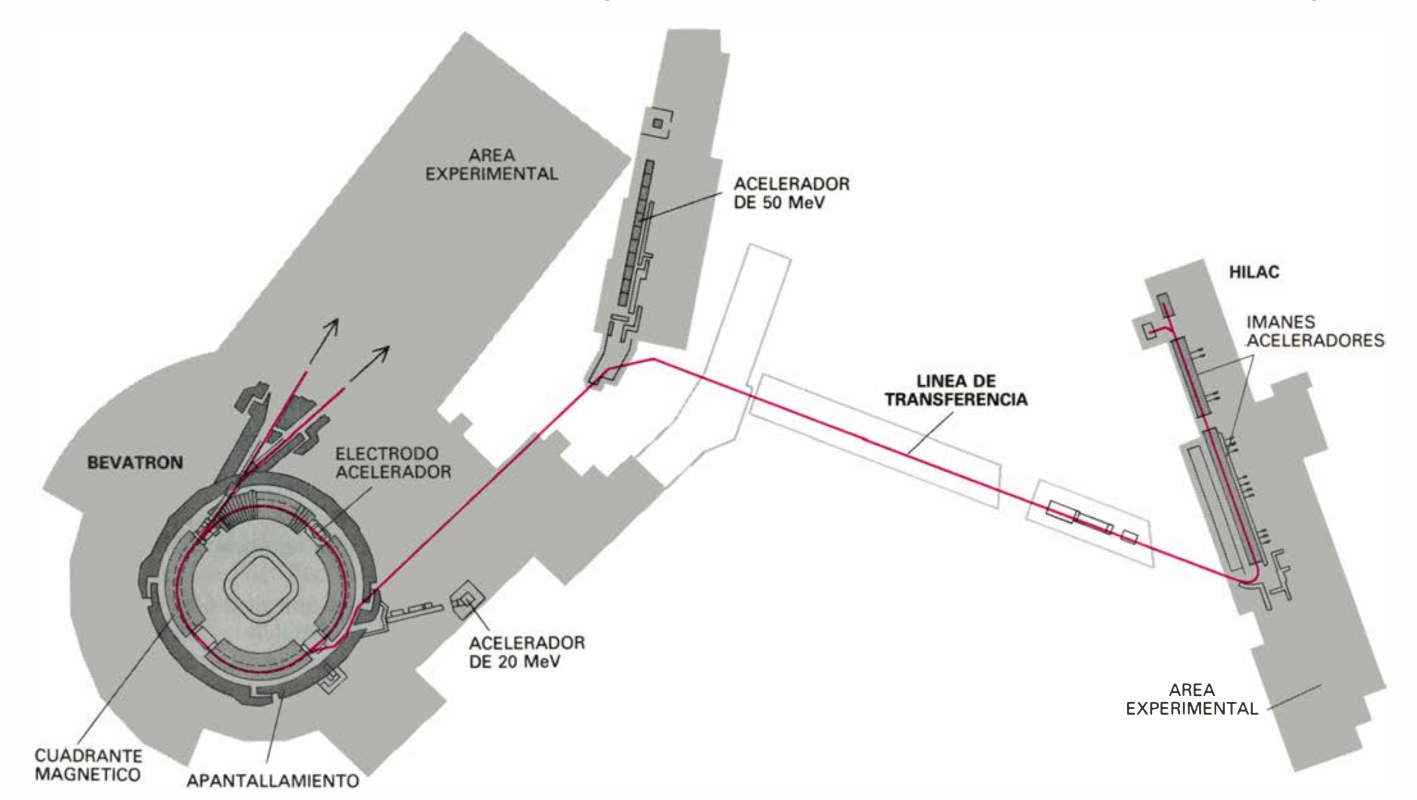
lugar a una traza de burbujas (en esencia, la partícula cargada provoca una ebullición local), y cámaras de dardos o cintas ("streamer"), en las que el paso de una partícula cargada por un gas, sobre el que se aplica un alto voltaje, deja una traza de dardos, o descargas eléctricas. Todos estos detectores confirman que los fragmentos de una colisión periférica llevan una dirección no muy distinta de la que porta el proyectil.

Estos detectores, por sí solos, no pueden dar el momento o la velocidad de los fragmentos. Para ello se pone en juego una técnica complementaria: se hace que la colisión ocurra en un campo magnético, que curva las trayectorias de los fragmentos, si llevan carga eléctrica. Cuanto mayor sea el momento de un fragmento, menor será la curvatura de su trayectoria; curvatura que se mide a partir del registro visual de la trayectoria o mediante una serie de detectores de partículas colocados en torno a la colisión. Las medidas muestran que los fragmentos de una colisión periférica poseen casi la misma velocidad que la que tenía el proyectil. Las pequeñas variaciones que se presenten pueden relacionarse con el tamaño del fragmento; los fragmentos de masa igual a la mitad de la del proyectil mostrarán una desviación máxima. La variación de la velocidad es del orden del cambio en velocidad que presenta un nucleón cuando se mueve por el núcleo. (Conocido por movimiento de Fermi, viene a ser aproximadamente el mismo en todos los núcleos.)

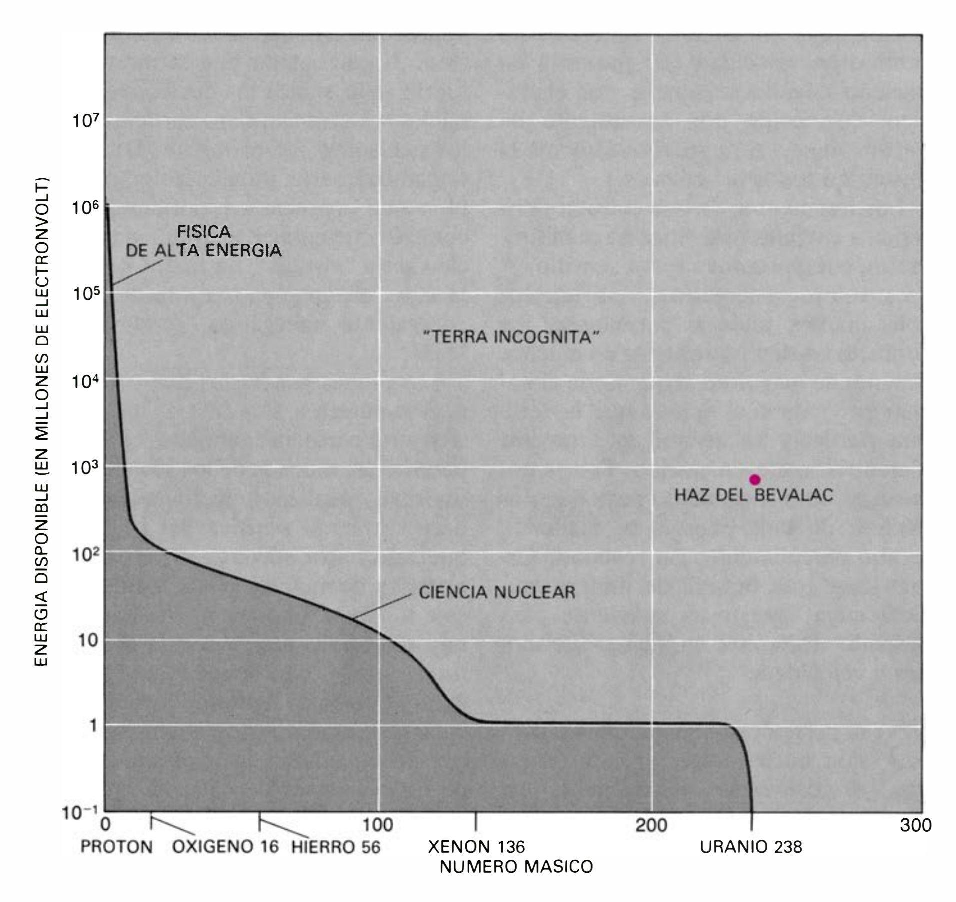
Los fragmentos de una colisión periférica a energías más bajas no manifiestan un comportamiento tan sencillo. A esos valores energéticos, los núcleos colisionantes pueden permanecer en contacto un tiempo superior en muchos órdenes de magnitud al que tarda el núcleo en vibrar o al tiempo que invierte una partícula relativista en atravesar diametralmente un núcleo. Por tanto, basta un leve rozamiento para que dos núcleos de baja energía se fusionen, total o parcialmente. En realidad, forman una gota líquida de materia nuclear cuya energía es suficiente para despedir nucleones en varias direcciones y velocidades.

Gran parte de la energía de una colisión núcleo-núcleo a altas energías se convierte normalmente en masa, a través de una copiosa producción de partículas subatómicas inestables y exóticas, principalmente piones, que son los mesones más ligeros. Estos se han visto implicados en una azarosa historia. En los años treinta, Hideki Yukawa predijo su existencia como "partículas virtuales" dentro del núcleo. Argumentaba que la interacción fuerte (que enlaza los nucleones en un núcleo) la transmitían partículas que los nucleones intercambian. Dicho intercambio sería indetectable; así, un pion que apareciera espontáneamente como fluctuación en el estado de un núcleo sería "virtual", no real. Con todo, la masa de un pion, expresada por su equivalente energético, sería de 139 MeV.

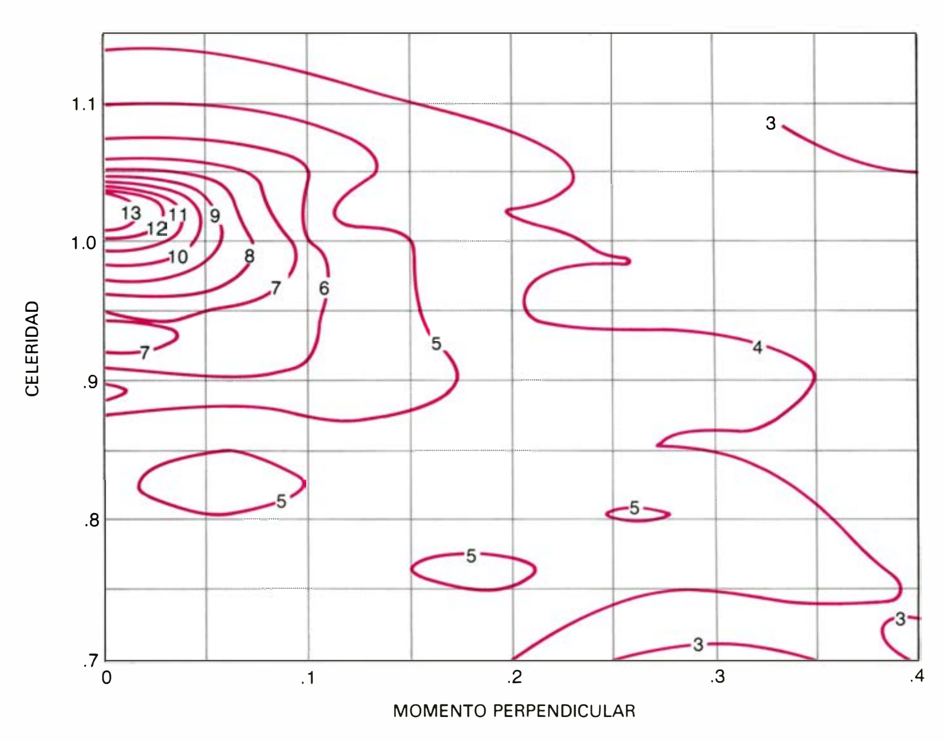
Los piones generados por colisiones núcleo-núcleo a alta energía ofrecen la siguiente peculiaridad: en las colisiones periféricas, muchos de los piones negativos se focalizan eléctricamente, es decir, la carga positiva del fragmento que sale hacia adelante atrae la carga negativa de cada pion, de manera que éste tiende a adquirir una velocidad y una dirección muy parecida a las del fragmento. Lo que puede tener aplicaciones prácticas. Aceleradores de partículas de Los Alamos, Vancouver y Zurich producen un haz de piones a partir de las colisiones de haces intensos de protones de alta energía; con un acelerador de iones pesados se podrían producir haces de piones sin tanto dispendio. Los iones pesados adecuados serían núcleos de silicio 28. Sucede que los piones y los núcleos de silicio 28, cuando atraviesan la materia, presen-



3. EL ACELERADOR DE NUCLEOS PESADOS de Berkeley consta de dos aceleradores conectados por una línea de transferencia del haz. El Bevatrón (*a la izquierda*), en forma de anillo, de 30 metros de diámetro, se construyó a comienzos de los cincuenta. Se proyectó para acelerar protones a 6 GeV. En los setenta resultaba ya anticuado; pero se vio que era capaz de acelerar núcleos pesados. Hoy, el Acelerador Lineal de Iones Pesados [HILAC], a la derecha, acelera núcleos a una energía de 8,5 millones de electronvolt (8,5 MeV) por nucleón (neutrón o protón). El Bevatrón los lleva hasta 2 GeV por nucleón y a velocidades de hasta el 95 por ciento de la velocidad de la luz. Se conoce por Bevalac la combinación de ambos aceleradores: Bevatrón e HILAC.



4. UNA NUEVA CIENCIA, explorada con las colisiones núcleo-núcleo a alta energía, se manifiesta cuando se representa la energía proporcionada a las partículas en el laboratorio (*eje vertical*) con respecto a su masa (*eje horizontal*). Las energías más altas se han impartido a las partículas de menor masa (subatómicas la mayoría de ellas), dejando una "terra incognita" que ocupa gran parte de la figura. La energía del haz de núcleos de uranio 238 del Bevalac (*punto rojo*) lo coloca en pleno territorio desconocido.



5. CONSTITUYE LA FOCALIZACION de piones un atributo de las colisiones núcleo-núcleo a alta energía; así, un haz de argón 40 de 21 GeV colisionando contra un blanco de cloruro potásico, en el Bevalac, produce piones negativos "focalizados" que tienen la velocidad del haz. Los fragmentos más pesados de las colisiones, que portan carga positiva, atraen los piones negativos. El gráfico representa la "rapidez", medida apropiada para velocidades próximas a la de la luz, en función del momento perpendicular a la dirección del haz. Las curvas de contorno muestran los números relativos de piones negativos.

tan un idéntico ritmo de pérdida de energía con respecto a la energía total. Resulta, pues, indiferente en qué punto de la materia ocurra la colisión que produce piones: el núcleo de silicio 28 que atraviesa un blanco puede chocar con un núcleo en cualquier punto interior al blanco; los piones negativos resultantes saldrán bien colimados en velocidad con los piones producidos en otras colisiones.

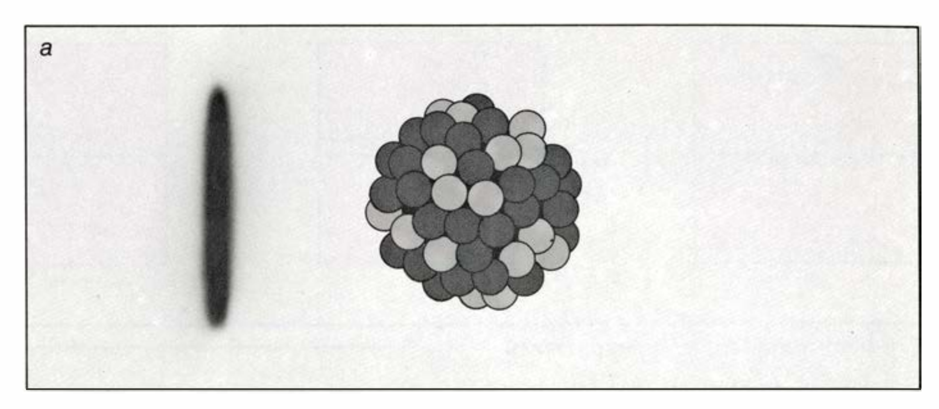
Hay un segundo producto de las colisiones periféricas núcleo-núcleo a alta energía (y de las posibles colisiones híbridas) que encierra mayor misterio. Nos referimos al anomalón: un fragmento de proyectil cuya trayectoria libre media (la distancia que recorre antes de chocar de nuevo) es corta en demasía. Quizá la mejor manera de describir el anomalón sea considerarlo núcleo que interacciona antes de tiempo. En los dos últimos años se han acumulado las pruebas de la existencia real del anomalón. Pero los experimentos son tediosos: exigen seguir series de colisiones, a través del examen de emulsiones fotográficas con un microscopio. La propia naturaleza de las pruebas recogidas, un pequeño número de interacciones presumiblemente prematuras, nos aconseja precaución. Los expertos en física nuclear se dividen hoy, a este propósito, en dos bandos: creyentes e incrédulos. ¿Por qué han despertado los anomalones tanta atención y controversia? De existir, quizá comporten implicaciones de largo alcance, no sólo para la ciencia nuclear sino también para la física de partículas elementales y, en consecuencia, para las teorías relativas a la estructura fundamental de la materia.

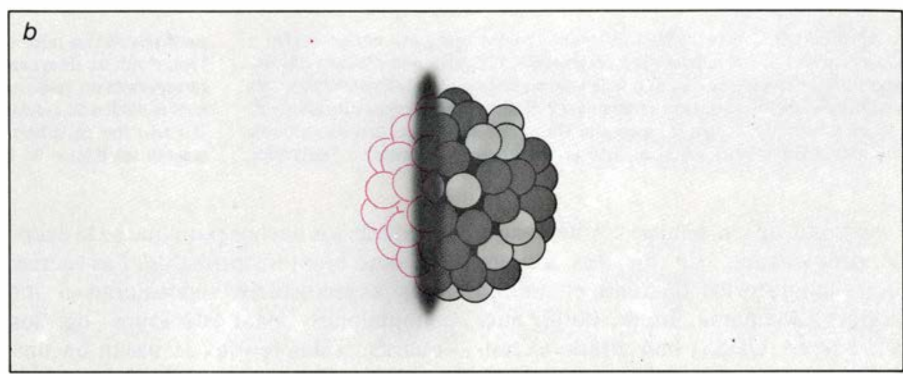
Los anomalones se observaron ya a principios de la década de los cincuenta, en los registros fotográficos de las colisiones de rayos cósmicos. De vez en cuando, un núcleo pesado de los rayos cósmicos chocaba contra un núcleo en la emulsión fotográfica, provocando una "estrella" (un estallido de trazas fotográficas); a continuación, un fragmento de la colisión creaba una segunda estrella a una distancia mucho más corta de lo predicho: quizás un centímetro o un par de ellos, frente a la decena o más acostumbrada.

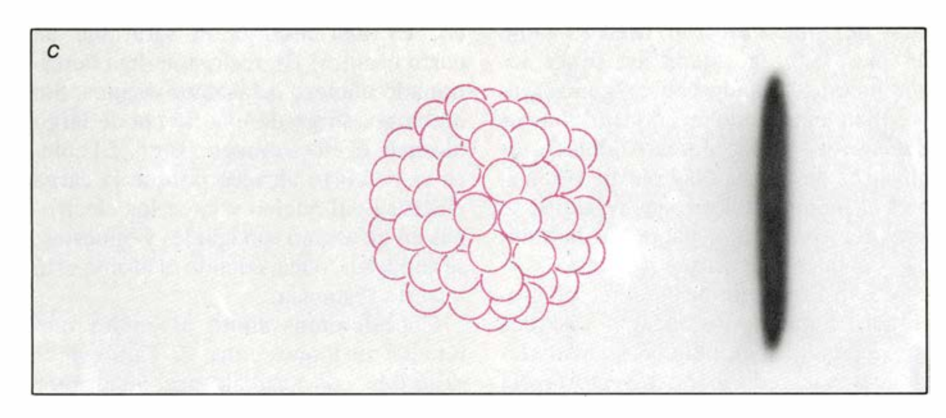
Cabe imaginar la perplejidad de los investigadores a la hora de explicar tan curioso suceso. Supongamos que reiteradamente hiciésemos el experimento consistente en conducir un automóvil por una autopista a gran velocidad y sacarlo de repente de la carretera hacia un bosque de grandes árboles distribuidos al azar. Unas veces nos estrellaría-

mos casi de inmediato; otras salvaríamos muchos árboles y, llegaríamos más lejos. Tras varias pruebas, habríamos establecido un recorrido medio, dependiente de la anchura del coche y el espaciado medio entre los árboles. En una serie posterior de pruebas usemos automóviles idénticos, con una sola diferencia: han sido reparados después de uno o varios choques anteriores. Quedaríamos asombrados si los viésemos chocar contra los árboles tras haber recorrido una distancia promedio diez veces menor que la recorrida por los coches nuevos. Evidentemente habría pasado algo imprevisto. Después de un choque, los coches se comportan como si fueran mucho más anchos. A tales coches los calificaríamos de anomalones.

Los registros fotográficos de las colisiones de rayos cósmicos recogieron esporádicas observaciones de anomalones a lo largo de los años cincuenta y sesenta. La aparición de los haces del Bevalac proporcionó registros de colisiones controladas núcleo-núcleo a alta energía en cantidad suficiente para calibrar las propiedades de los anomalones. Tres grupos, uno del Lawrence Berkeley Laboratory, otro del National Research Council de Canadá, en Ottawa, y un tercero de la Universidad estatal de Nueva York en Buffalo, examinaron emulsiones fotográficas expuestas a los haces de varios núcleos y varias energías del Bevalac, llegando hasta núcleos de hierro 56, a 1,88 GeV por nucleón. Otro grupo de la Universidad de Minnesota volvió a examinar los registros de rayos cósmicos. Cuando se celebró en Berkeley, hace dos años, el primer seminario internacional sobre anomalones, todos esos esfuerzos conjuntos condujeron a los resultados que vamos a detallar.



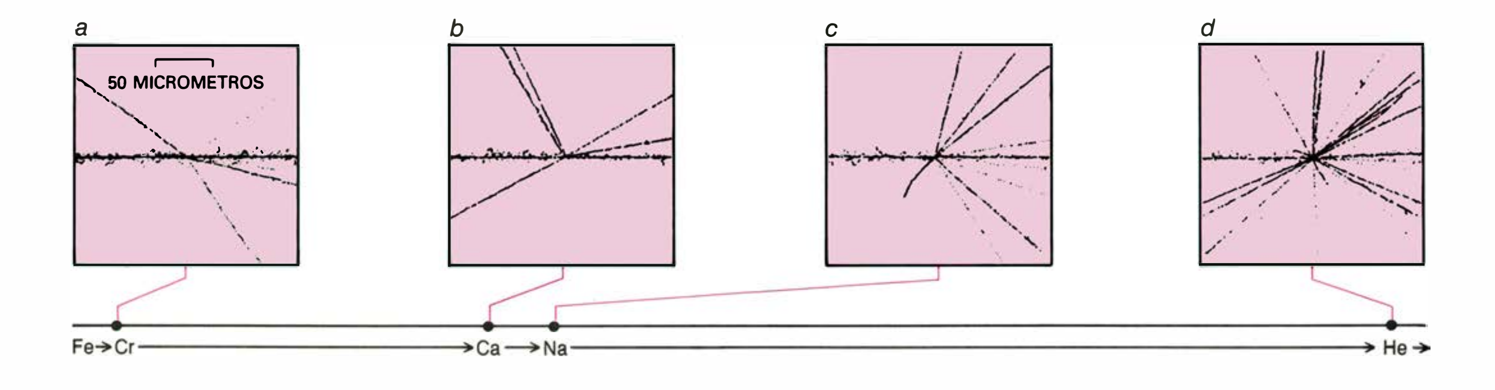




6. CONTRACCION RELATIVISTA, otro atributo de las colisiones. Se ilustra una hipotética colisión a energía ultra-alta. Un núcleo de uranio, con una energía de un billón de electronvolt por nucleón, se aproxima a un núcleo de uranio en reposo (*a*). El proyectil, que viaja a más del 99,999 por ciento de la velocidad de la luz, aparece en forma de disco; su contracción es predicha por la teoría especial de la relatividad. Su encuentro con el blanco sólo dura 10^{-22} segundos, tiempo insuficiente para que los dos alcancen un equilibrio; así, el proyectil pasa de largo (*b*). Pero eleva el blanco a una temperatura de más de un billón de grados (*c*), la que había en el universo inmediatamente después de la gran explosión.

En primer lugar, los anomalones tenían recorridos libres medios de desusada brevedad. Entre los fragmentos del proyectil examinados por los distintos grupos, un pequeño porcentaje (en torno al 6 por ciento, aunque las cifras presentadas variaban del 3 al 10 por ciento) alcanzaba recorridos libres medios no mayores que 2,5 centímetros, menos de la décima parte de la longitud de las trayectorias de los proyectiles. Segundo, los anomalones sólo se dejaban ver cuando la energía de los proyectiles superaba el GeV por nucleón (se habían hecho pocos trabajos por debajo de 1,8 GeV por nucleón, de modo que se desconocía el umbral de energía preciso). Tercero, los anomalones debían tener una vida media del orden de 10^{-10} segundos, por lo menos; de lo contrario, no habrían recorrido 2,5 centímetros. Tal vida media es, aproximadamente, 13 órdenes de magnitud mayor que el tiempo típico de sucesos tales como la vibración de un núcleo. Cuarto, si los anomalones se desintegraran espontáneamente en vuelo, la desintegración debería estar acompañada sólo por la emisión de partículas de carga neutra, a saber, neutrones o rayos gamma (cuantos de alta energía de radiación electromagnética), puesto que no se observaron partículas cargadas que emanaran de las trayectorias de los anomalones a través de las emulsiones fotográficas. Quinto, las propiedades de los anomalones parecían independientes de la cantidad de carga del fragmento. Alguna leve indicación sugería que nunca se formaban anomalones con sólo una o dos unidades de carga.

Sexto, y quizá lo más extraño de todo, los anomalones presentaban el siguiente efecto de memoria: “anomalón una vez, anomalón siempre”. Con mayor precisión, parece ser que un fragmento que emerja de una estrella producida por un anomalón tiene una probabilidad mayor de constituir, a su vez, un anomalón que la que tendría si emergiera de una estrella producida por algo distinto. Así pues, o los anomalones tienen buenas probabilidades de sobrevivir tras una colisión o algo relacionado con los anomalones hace que sea relativamente fácil reconstruirlos en una colisión.



7. ANOMALON, peculiaridad intrigante de las colisiones núcleo-núcleo a alta energía. Los anomalones son fragmentos nucleares que realizan ulteriores colisiones dentro de una distancia anormalmente corta. (Se observaron ya a principios de los años cincuenta en las colisiones de los rayos cósmicos.) El esquema superior ilustra un ejemplo. Un núcleo de hierro, con una energía de 1,88 GeV por nucleón, chocó con un núcleo en una emulsión fotográfica, produciendo un núcleo de cromo (*a*). Después de un vuelo de 1,62 centímetros, el núcleo de cromo choca, produciendo un núcleo de calcio (*b*); se hubiera esperado un vuelo medio de 10 centímetros o más. Después de 0,3 centímetros el núcleo de calcio colisiona, originando un núcleo de sodio (*c*); de nuevo el vuelo fue cortísimo. Tras recorrer 3,75 cm, el núcleo de sodio choca y origina un núcleo de helio (*d*). Las colisiones se registraron en el Bevalac.

Se trata de un conjunto interesante de propiedades. En los dos últimos años, laboratorios de todo el mundo (Egipto, Alemania, India, Italia, Suecia, Suiza y URSS) han estado examinando registros fotográficos del Bevalac y del Sincrofasotrón. Merced a ello ha mejorado la estadística sobre los anomalones. Añádese que algunos grupos han empezado a emplear nuevos detectores, especialmente detectores plásticos de trazas. Hay plásticos sensibles al ataque de los iones pesados; se exponen, por tanto, paquetes de láminas de tales plásticos a los iones y se procesan luego químicamente. El procesado aumenta los lugares dañados convirtiéndolos en pequeñas cavidades cónicas que indican la trayectoria del ion. Esta técnica ofrece una gran ventaja: a cierta velocidad, los iones producen una cavidad cuyo diámetro mide su carga; el diámetro de la cavidad puede determinarse mediante escrutadores computarizados. Cuando el verano pasado se celebró, en Berkeley, el segundo seminario internacional sobre anomalones, los participantes reconocieron la validez de esa lista de seis propiedades, aunque admitían la falta de pruebas determinantes y varios grupos informaron que la carga eléctrica de los anomalones parecía ser siempre un múltiplo de la carga eléctrica unidad. De donde se introducía una séptima propiedad: los anomalones carecían de cargas eléctricas fraccionarias. En otras palabras, los anomalones no parecen incluir quarks aislados.

Adivinar una explicación teórica satisfactoria de los anomalones es, por lo menos, tan difícil y complicado como lo ha sido el trabajo experimental. Se han propuesto muchas teorías, pero hasta ahora ninguna ha dado suficiente razón de los hechos para que se la acepte como hipótesis probable. Las teorías más espectaculares consideran a los anomalones manifestaciones de los quarks. Tales teorías se basan en una analogía. Pensemos en el enlace químico. Es una interacción saturable de corto alcance. Es decir, sólo liga determinado número de átomos vecinos. Sin embargo, surge de una fuerza de largo alcance: el electromagnetismo. El enlace es de corto alcance porque la carga eléctrica del núcleo y la de los electrones en un átomo son iguales y opuestas; se cancelan, pues, cuando el átomo está a cierta distancia.

Consideremos ahora el enlace que une los nucleones entre sí. También es saturable y de corto alcance. Su alcance normal es menor que el tamaño de un núcleo. Pero puede también que surja de una fuerza de largo alcance. El enlace que liga los nucleones entre sí sería entonces de corto alcance por la exclusiva razón de que cada uno de los tres quarks de un nucleón tiene una carga de fuerza fuerte, llamada "color", y los colores normalmente se cancelan.

Los anomalones proporcionan una de las primeras bases reales para pensar en una posible fuga de largo alcance en la fuerza fuerte. De ahí los esfuerzos realizados para visualizar los anomalones cual reuniones de quarks donde no se cancelan los colores, de modo que una interacción fuerte de largo alcance facilita las colisiones y reduce el recorrido libre medio del fragmento nuclear. Los intentos han ido desde el "deuterón diabólico" (un deuterón formado por tres pares de quarks, en vez de dos tripletes de ellos) hasta grandes masas de quarks en forma de rosquilla. Ningún intento ha resultado compatible con todas las propiedades de la lista. El hecho de que los anomalones se creen a energías menores que las que uno esperaría para una importante redistribución de los quarks, o para una copiosa producción de los mismos, no es, precisamente, una dificultad menor.

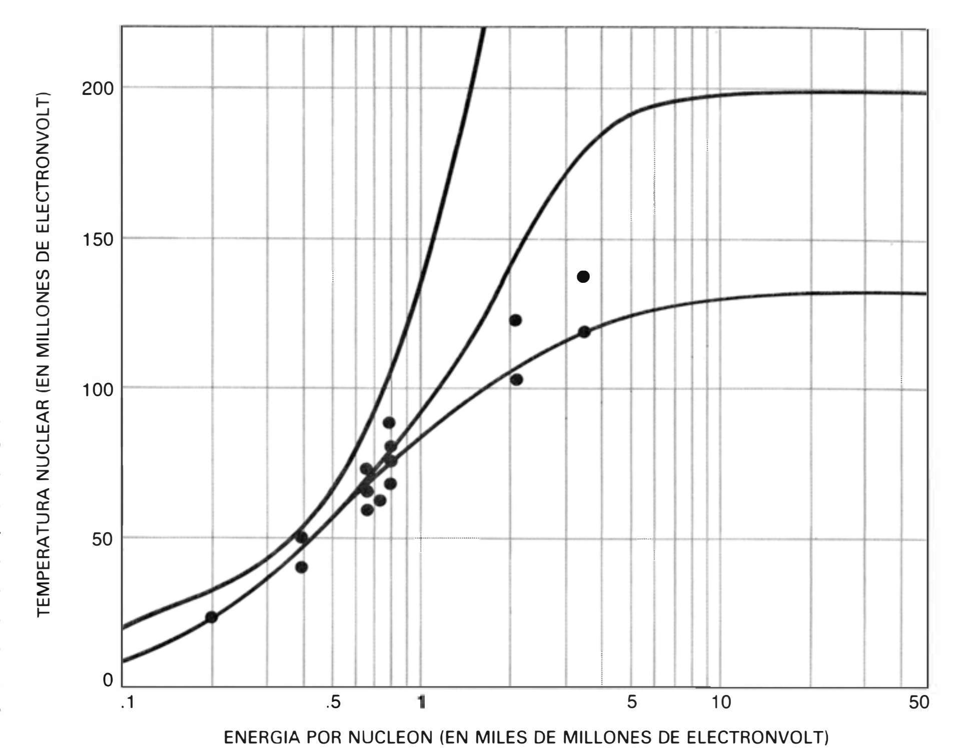
Las explicaciones de los anomalones que no se basan en los quarks han sido también diversas. Van desde núcleos en burbuja (pelotas huecas de nucleones) hasta núcleos cuasi-moleculares (agregados en forma de pesas de gimnasta que se obtienen cuando dos núcleos se unen pero no se amalgaman). De lo que se trata es de convertir el anomalón en un objeto hinchado o distendido, para que aumente la probabilidad de ser golpeado en una colisión. Entre las explicaciones se cuenta la nuestra: el anomalón consta de piones negativos ligados fuertemente a fragmentos nucleares ricos en neutrones. Los piones podrían dirigir, por convección, los neutrones hacia fuera de la materia nuclear, y extender así la superficie, haciendo que el fragmento presente protuberancias. En nuestra opinión, los piones se producen con abundancia en las colisiones núcleo-núcleo y los negativos se focalizan en torno a los fragmentos mayores. Además debe recordarse que los piones son atraídos hacia los protones y neutrones por la interacción fuerte. Y se sabe que la atracción de los piones negativos es mucho mayor hacia los neutrones que hacia los protones.

Nuestro modelo topa con una dificultad: cualquier contacto de los piones negativos con los protones haría que la materia se desintegrara en tiempos cortos, comparado con la vida media de los anomalones. Sin embargo, nos anima el que cálculos teóricos recientes sugieran que un cuerpo que conste de algunos piones negativos y otros pocos neutrones alcance una vida lo bastante

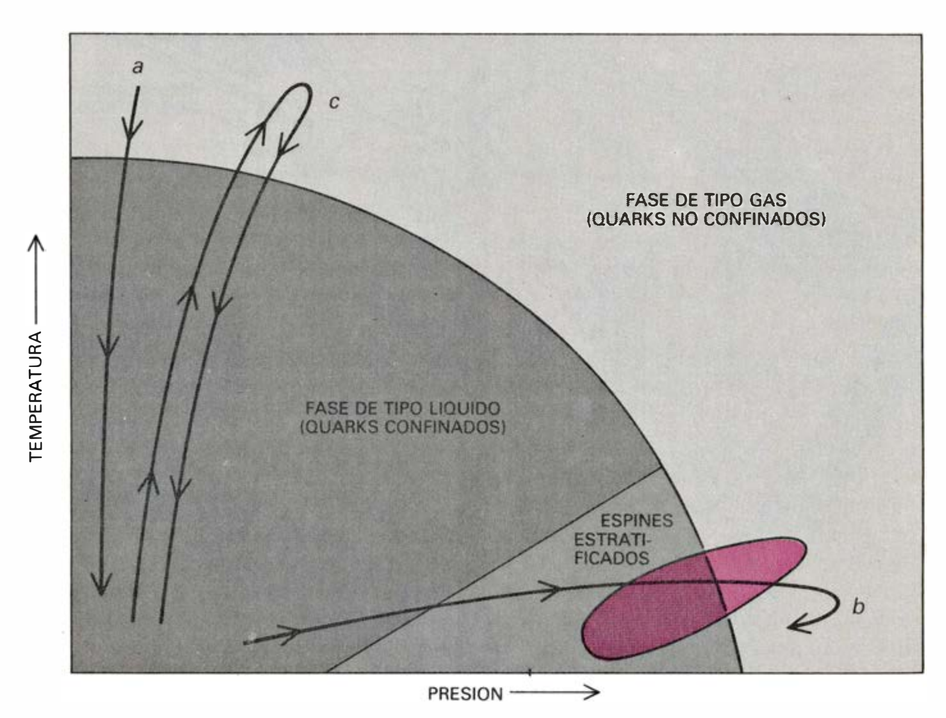
larga para detectarlo. Nadie se ha encontrado nunca con ninguna bola de piones de ésas, pero su búsqueda en varios aceleradores de iones pesados puede quizá fructificar. De hecho, está empezando a verse que los piones negativos ligados a núcleos constituyen un fenómeno no exclusivo de una teoría de anomalones. Algunos datos recogidos por los detectores de partículas indican que los piones han orbitado en torno a fragmentos nucleares para salir después despedidos. En consecuencia, entendemos que el anomalón vendría a ser una bola de piñones rodeando un centro nuclear, con la fuerza centrífuga manteniendo los piñones alejados de los protones del centro. Quizás el necesario núcleo rico en neutrones (o la región del mismo rica en neutrones) explica el efecto de memoria que se observa en los anomalones. Un fragmento nuclear rico en neutrones tendería, en sucesivas colisiones, a producir más fragmentos ricos también en neutrones.

¿Puede establecerse una conexión entre piones negativos ligados a neutrones y una fuga de largo alcance de la fuerza fuerte? Hasta ahora, los ensayos para abordar los núcleos sencillos como agregados de quarks se han centrado en múltiplos de tripletes de quarks. Tales múltiplos garantizan virtualmente la cancelación de las cargas de color. ¿Qué hay acerca de los cálculos basados en un pion negativo y dos neutrones, por ejemplo? Esta es la bola de piñones más sencilla. Constaría de ocho quarks: cinco quarks abajo, dos quarks arriba y un antiquark abajo (un pion negativo se supone formado por una pareja quark abajo-antiquark arriba). Parece menos probable que tal agregado garantice la cancelación de las cargas de color; su estudio debe continuar.

Las colisiones núcleo-núcleo a alta energía presentan una peculiaridad particularmente atractiva para la investigación cosmológica: el impresionante ritmo con el que un núcleo pesado de alta energía deposita energía sobre un núcleo blanco sugiere que este núcleo blanco podría calentarse a una temperatura cercana a la del universo al poco de la gran explosión. Tales temperaturas no podrían alcanzarse bombardeando núcleos pesados con haces de electrones o protones; la levedad de estos proyectiles los haría disipar sólo una pequeña fracción de su energía cinética total cuando atravesaran hasta el más pesado de los núcleos. En cambio, una colisión central entre dos átomos de uranio podría detener la marcha del proyectil, que descargaría así toda su energía cinética.



8. TEMPERATURAS NUCLEARES alcanzadas en colisiones núcleo-núcleo a energías del Bevalac. Sirven de banco de pruebas de distintos modelos sobre materia nuclear en condiciones extremas. Las curvas son cálculos teóricos que involucran diferentes hipótesis acerca de la materia nuclear. De acuerdo con una de éstas (*curva inferior*), el calentamiento de los núcleos convierte toda la energía disponible en materia, es decir, en hiperones, estados excitados de los nucleones. Los hiperones pueden desintegrarse después en nucleones y piones. La hipótesis opuesta (*curva superior*) defiende que el calentamiento tiene lugar sin modificar los nucleones en otras partículas. Una tercera hipótesis (*curva central*) es intermedia. En el esquema, todas las temperaturas están expresadas por su equivalente en energía cinética en MeV.

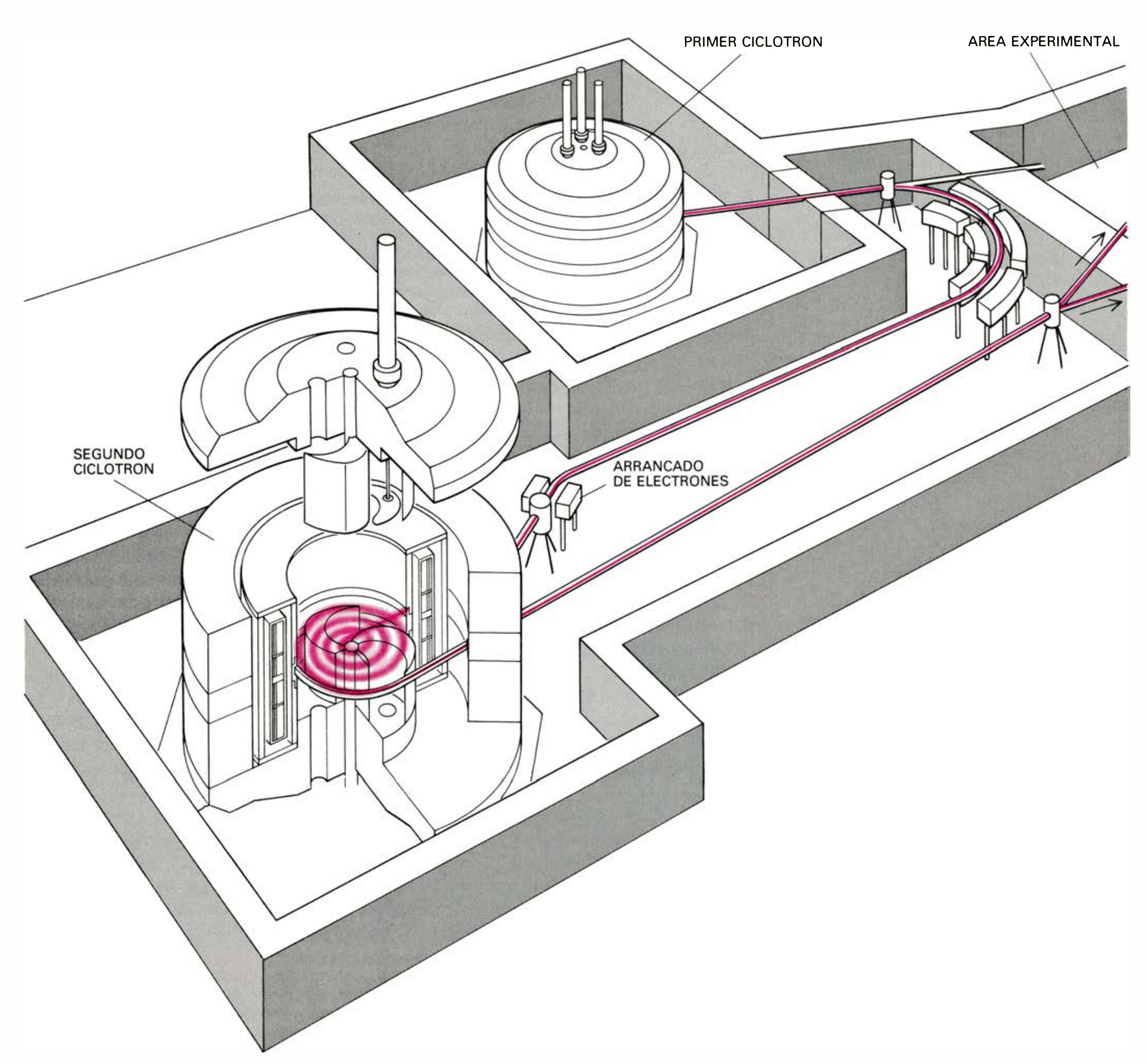


9. DIAGRAMA DE FASES HIPOTETICO, donde se ofrecen las formas de materia nuclear que pueden surgir bajo condiciones, ampliamente variables, de temperatura y presión. Los núcleos normales están en una fase parecida a la líquida (*gris oscuro*): son un "fluido" de protones y neutrones. A temperaturas y presiones elevadísimas, estos nucleones se disocian en quarks y gluones, sus supuestos constituyentes (*gris claro*). A combinaciones de alta presión y temperatura intermedia existe, quizás, una tercera fase, donde los espines de los nucleones se organizan en estratos (*gris medio*). Una trayectoria del gráfico muestra la evolución de la materia nuclear en el universo primitivo (*a*). Una segunda trayectoria sigue la evolución de la materia nuclear en la explosión de una supernova (*b*). Y una tercera repite la evolución de la materia nuclear en una colisión central núcleo-núcleo (*c*). Tal vez puede conseguirse un reino diferente a través de las colisiones núcleo-núcleo periféricas, el que recoge el óvalo en color (Ilustración de Walken Graphics.)

Un equipo germano-estadounidense ha examinado en el Bevalac hasta qué punto los núcleos a máximas energías de dicha máquina se paran realmente en las colisiones centrales. Han construido una capa esférica muy elaborada de detectores múltiples que determinan, a la vez, energía y ritmo de deposición de ésta en las partículas que salen del blanco. Unos pocos metros más en la dirección del haz, hay otro sistema de detectores. El conjunto total, "pelota" y "pared", incorpora más de mil detectores. Los datos muestran que los núcleos de calcio son demasiado pequeños para parar uno a otro; no así los núcleos de niobio.

La noción de temperatura del núcleo exige una explicación. En general, la temperatura es simplemente una medida de los movimientos aleatorios en un grupo de partículas. La energía de una colisión nuclear se transforma, al menos en parte, en movimientos aleatorios de los nucleones del núcleo; por tanto, es apropiado hablar de una temperatura nuclear. De manera específica, es característico de la materia en equilibrio térmico, conste de nucleones, de moléculas, o de partículas macroscópicas en movimiento aleatorio, que la distribución de velocidades disminuya exponencialmente. En otras palabras, los constituyentes de alta energía son relativamente escasos. La energía cinética media es proporcional a la temperatura, T, de la materia, expresada en grados Kelvin; la constante de proporcionalidad es la constante de Boltzmann, k. Los científicos nucleares prefieren no trabajar con T, que puede tomar valores excesivos (miles de millones de grados). Emplean como sustituto kT, la energía cinética media por nucleón.

Demos algunos ejemplos. La temperatura de la superficie del Sol (T) es de unos 6000 grados K. La correspondiente energía cinética kT es 0,5 electronvolt. Ello significa que los átomos de la superficie del Sol tienen menos energía



10. CICLOTRONES SUPERCONDUCTORES instalados en la Universidad estatal de Michigan. Los primeros de su serie. Un par de ellos acelerarán núcleos pesados a energías de hasta 200 MeV por nucleón. Un ciclotrón, ya acabado, acelerará átomos ligeramente ionizados a lo largo de una trayectoria espiral entre los polos helicoidales de un imán superconductor de sólo 3 metros de diámetro. (La región de aceleración propiamente dicha tiene sólo 132 centímetros de diámetro.) El haz de iones resultante será dirigido, a través de una lámina que despoja a los átomos de muchos de los electrones que aún tengan, hacia el segundo ciclotrón, que acelerará todavía más a los iones. Este par de máquinas resultarán de tamaño y costo modestos, si lo comparamos con los aceleradores recientes, pero generará haces cuya energía por nucleón superará la conseguida por la mayoría de aceleradores de iones pesados.

cinética que los electrones acelerados por el voltaje de la batería de una linterna. (Obsérvese que la energía cinética adquirida por una partícula al calentarse no suele ser enorme.) Se cree que la temperatura en el centro del Sol es de 15 millones de grados. La energía cinética correspondiente es de 1300 electronvolt, es decir, aproximadamente la energía de los electrones que inciden sobre la pantalla de un televisor. La temperatura en el centro de una estrella cuando se convierte en una supernova podría ser de unos pocos miles de millones de grados. La energía cinética correspondiente alcanza 1 MeV, es decir, la energía impartida a los electrones en los más potentes microscopios electrónicos. Este es el límite superior de las temperaturas generadas en procesos que se desarrollan en el universo actual. Se cree que hubo temperaturas mucho mayores en el universo primitivo.

Cabe preguntarse si sucede algo especial a la materia nuclear cuando se lleva a temperaturas extremadamente altas. Consideremos qué le ocurre a la materia ordinaria sometida a calentamiento. Partamos de una materia en fase sólida. A medida que se calienta, funde, convirtiéndose en líquido; luego hierve, convirtiéndose en gas. Sus moléculas se descomponen después en átomos. Finalmente, los electrones se separan de los átomos; electrones y núcleos ionizados, algunos de ellos completamente despojados de electrones, se mueven independientemente en un gas de partículas cargadas: un plasma.

Se cree que la materia nuclear no se halla en fase sólida, sino que se parece a un líquido. Sin embargo, los físicos teóricos predicen que una alta temperatura (o una gran presión) llevaría a una transición de fase, en la cual los nucleones se romperían en quarks y gluones: las partículas que los quarks intercambian para unirse entre sí. El resultado sería algo semejante a un plasma: los quarks se moverían por todo el núcleo, no sólo dentro de la "bolsa" del triplete de los mismos.

Esta caracterización de la materia nuclear se mueve, obviamente, en un plano especulativo. Podría defenderse incluso una fase de "condensado de piones", a altas densidades nucleares sin altas temperaturas, amén de la fase normal, líquida, y la predicha de plasma. En esa nueva fase, los nucleones se ordenarían en estratos, de manera que el espín, o momento angular intrínseco, de los nucleones de un estrato estaría alineado en una dirección y el espín del estrato siguiente se alinearía en la dirección opuesta. ("Condensación de piones" no es quizá la denominación más idónea; tiene que ver con la resultante estabilidad del enlace de piones en materia nuclear densa.) Se ha llegado incluso a suponer que en el centro de las estrellas de neutrones habría materia nuclear en la fase de condensado de piones. Por desgracia, los esfuerzos para producir tal fase en la Tierra se enfrentan con el problema de dar a la materia nuclear una densidad elevada sin elevar su temperatura.

¿Cuán caliente y densa debe ser la materia nuclear para convertirse en un plasma de quarks y gluones? Aportar una respuesta con cierta credibilidad exige un trabajo que mantiene ocupados potentes ordenadores. Aunque no se dispone de resultados definitivos, las mejores estimaciones sugieren que la temperatura, a densidades nucleares normales, debería aproximarse a 200 MeV. Tales condiciones están, probablemente, fuera del alcance de la generación actual de aceleradores de iones pesados. Pero se están preparando experimentos para buscar pruebas por si pudiera haberse creado ya un plasma de quarks y gluones.

Nuevos aceleradores proporcionarán mejores posibilidades. Hay algunos terminados y otros en construcción. Gracias a los imanes superconductores y a fiables controles microelectrónicos, pueden hacerse más compactos y menos costosos que lo que sus propios defensores habían temido. Sin embargo, los haces nucleares de energías altísimas, necesarios para la exploración de las fases nucleares y sus implicaciones cosmológicas, pertenecen todavía al futuro. El Lawrence Berkeley Laboratory y el Brookhaven National Laboratory han adelantado proyectos provisionales de aceleradores de iones pesados (con los asociados anillos de almacenamiento de iones) en el rango de los TeV (billones de electronvolt).

Los aceleradores disponibles dan una ligera idea de la ciencia que trata de las colisiones relativistas núcleo-núcleo. Se trata de una ciencia híbrida, basada en conocimientos de dos campos: física de partículas subatómicas de alta energía, por un lado, y ciencia nuclear, por otro. En las últimas décadas, ambas han ido distanciándose y tienen hoy día poco en común. Pero al igual que ocurre con la unión de dos climas, que no sólo genera una gran turbulencia, sino también copiosa fertilidad, así la unión de ambas ciencias, física de partículas y ciencia nuclear, dotará de extraordinario vigor el estudio de las colisiones núcleo-núcleo.

Ciencia y sociedad

Física aristotélica

Se reconoce, en la *Física* de Aristóteles, una obra capital de la historia del pensamiento. Pero ya Martin Heidegger nos avisaba de que: "La *Física* de Aristóteles es el oculto, y por esto nunca suficientemente repensado, libro básico de la filosofía occidental". Vale decir: se encuentra hoy todavía necesitada de estudio.

Varios autores consideran, por otra parte, que tal estudio resulta especialmente congruente con la situación actual de la ciencia y con los profundos cambios que ha sufrido en particular la física, no sólo en cuanto a sus teorías generales, sino también en cuanto a las cuestiones de fundamentación y de status teórico de la misma. Lo que a más de uno le parecerá paradójico si piensa en un Aristóteles adversario a combatir por los pioneros de la llamada física clásica y personificación tópica del pensamiento acientífico. La verdad es que Aristóteles debe ser contado entre los impulsores, en general, del rigor intelectual, y en particular, del interés por la naturaleza. Y aunque en muchos dominios sus explicaciones ya no tengan otro interés que el histórico, su filosofía de la naturaleza (el contenido principal de la *Física*) conserva aún hoy un verdadero aliciente teórico.

Frecuentemente no se acierta al establecer la comparación entre nuestro modo de conocer la naturaleza y el aristotélico. Las controvertidas diferencias en torno al geocentrismo, la ley de inercia, la caída de los graves o la teleología, aunque han tenido un papel histórico eminente, no dejan de ser accidentales. La diferencia fundamental, la que más nos puede interesar, entre Aristóteles y hoy, es su distinto punto de partida: su distinta definición del ser natural. Lo natural, para nosotros, es lo situado en alguna parte del tiempo y del espacio, sea que se le entienda como cuerpo o como acontecimiento. Este modo de concebir es el propio de las ciencias experimentales, que, por las características de su método, no pueden determinar de otra manera a su objeto. Pero no es esa la noción primera de ente natural en Aristóteles, y, para los estudios del muismo, resulta una restricción definirlo así. Sin embargo, desde Descartes el pensamiento moderno se ha apegado innecesariamente a dicha definición, y ha limitado con ello nuestra comprensión del ser natural.

La concepción aristotélica entiende el ser natural como ente móvil, es decir, cambiante. Al considerar que lo que define el ser natural es la movilidad y el devenir mismo, no se cancela un ulterior tratamiento experimental-matemático, y se gana en las primeras nociones fundamentales sobre la naturaleza una mayor proximidad al ser, lo que conlleva, en dicho contexto, una inteligibilidad superior. A partir de la movilidad, Aristóteles derivará la necesaria corporeidad del ente natural, su espacio-temporalidad, el aspecto cuantitativo que esto comporta y las características de las magnitudes físicas como continuos. Ese aspecto del ente natural que funda la posibilidad de la determinación matemática de los procesos naturales no es, por tanto, según Aristóteles, lo primero y esencial del ente natural, sino una característica particular, derivada de lo esencial: el carácter cinético de su ser.

Aristóteles empieza su *Física* suponiendo, como un hecho primario y evidente que no necesita demostración, la existencia del movimiento en su sentido más genérico de cambio. Con ello cumple el primer requisito que, según sus propios estudios lógicos, toda ciencia en forma debe cumplir: dar por supuesta desde el comienzo la existencia de su objeto –en este caso, el ente móvil–. Este supuesto de la existencia del objeto es necesario para que la ciencia tenga carácter real, y no meramente verbal o formal. Acto seguido, el objeto debe ser definido; pues, en su opinión, la definición es el punto de partida de las demostraciones, por lo mismo que de la esencia fluyen las propiedades del objeto.

¿Dónde define Aristóteles con propiedad el objeto de su *Física*? Muchos autores se inclinan por el libro II, capítulo 1.º; pero si así fuera no se entendería la función del libro I. En éste se contiene la exposición fundamental del hilemorfismo, pieza clave de todo el pensamiento aristotélico; allí se describe la constitución esencial del ente móvil, lo que equivale a su definición: su ser compuesto de materia y forma, por decirlo con términos suyos.

A través de un análisis del lenguaje con el que expresamos el cambio, examina lo que concurre en el cambio mismo. En todo cambio, dice, hay que distinguir "lo que cambia" y "lo que viene a ser" con el cambio. En efecto, cambio implica siempre algún tipo de novedad, algo que se adquiere. Aristóteles lo llama "forma", tomando este término en el sentido amplio de "determinación de ser". Pero debe también haber algo que sufre el cambio y para lo cual la forma es una adquisición; pues sin ello no habría tampoco adquisición ni novedad: no hay cambio si nada cambia. Pues bien, "lo que cambia" puede considerarse de un modo doble, o bien como el contrario de la forma, o bien como su sujeto. Por ejemplo, si nos preguntamos ¿qué es lo que llega a ser culto? Podemos responder "lo no culto", o bien "el hombre". Ambos modos son necesarios, pues el cambio debe proceder de lo contrario, de la privación de la forma, pero, además, "lo que cambia" no puede ser sólo lo contrario. Es necesario que sea también algo más, algo apto para tener la forma y que, en vez de dejar de ser con el cambio, permanecerá a través de él. Como consecuencia, Aristóteles afirma que todo "lo que cambia", es decir, todo ente móvil, está compuesto necesariamente por un sujeto y una forma.

Cuando el cambio es absoluto se engendra o se destruye una sustancia. Por consiguiente, el sujeto no puede ser la sustancia. El pondrá al sujeto en un principio del ser incognoscible en sí y sólo pensable por analogía, totalmente informe y potencial: la materia aristotélica. La composición de materia y forma es, pues, la "composición sustancial" universal de todo ente móvil, y es lo que hace posible el cambio. (Con ello evita la aporía de los filósofos eleatas, que negaban el movimiento. A partir de este momento será posible la física.) A este modo de definir el ente móvil por sus principios constitutivos sólo le falta darle un giro preciso para que quede manifiesto el sentido en que se denomina también ente *natural*.

En el libro II, Aristóteles define el ente natural como aquel que posee en sí mismo un principio de movimiento y reposo; se llama "naturaleza" a ese principio intrínseco del movimiento. Puesto que, como acabamos de ver, la materia y la forma son los dos principios constitutivos del ente móvil en cuanto móvil, se desprende que a ambos les conviene el nombre de "naturaleza", si bien a la forma con mayor propiedad que a la materia.

Para Aristóteles, lo fundamental es que ese comportamiento no es algo puramente extrínseco, que se halle como de tránsito en las cosas, sino que tiene en ellas, en lo que son, su raíz y causa

principal. En este carácter intrínseco del principio del movimiento, que define lo natural como tal, reside la concepción aristotélica de la naturaleza. (Antonio Prevosti.)

... *Y astrofísica actual*

El último premio Nobel de física lo han compartido dos astrofísicos, Subrahmanyan Chandrasekhar, de la Universidad de Chicago, y William A. Fowler, del Instituto de Tecnología de California. Méritos del primero: desarrollo teórico y experimental de las reacciones nucleares implicadas en la estructura y evolución de las estrellas; méritos de Fowler: desarrollo teórico y experimental de las reacciones nucleares que producen la formación de los elementos químicos del universo. Se llama astrofísica la ciencia que se ocupa de las propiedades y estructura de las estrellas y otros cuerpos celestes. Nos explica qué procesos acontecen en los mismos, de dónde sacan la energía necesaria para las reacciones que experimentan, la evolución que siguen y el origen de los elementos, o nucleogénesis.

William A. Fowler se adentró en el estudio de las reacciones nucleares hace treinta años largos, trabajando en el ciclo del carbono-nitrógeno, es decir, en la secuencia de reacciones que producen la síntesis de helio a partir de hidrógeno y en la que intervienen, como catalizadores, el carbono, el nitrógeno y el oxígeno. (Se apoyaba en un célebre trabajo de Hans A. Bethe sobre dicho ciclo, de 1938.) A ese nuevo campo llegó Fowler con un buen bagaje experimental, asociado a Charles Lauritsen. Habían estudiado la captura de protones por elementos ligeros, producida en resonancias abruptas. Como ha escrito el propio Bethe en *Science*, a propósito del nuevo Nobel: "Era el planteamiento adecuado para investigar las reacciones del ciclo del carbono, de las que destaca, por encima de todas, la captura de protones con la emisión de rayos gamma. Fowler y sus colaboradores descubrieron que la tasa de las reacciones que se suceden en el ciclo del carbono-nitrógeno era la idónea para justificar la producción de energía en las estrellas dotadas de gran masa de la secuencia principal". Fowler prefiere hablar de biciclo del carbono, ya que el isótopo O-17 –formado por captura de un protón y emisión de un antielectrón– reacciona con el H para rendir N y He-4.

En las estrellas hay que distinguir el núcleo de la superficie; aquél muy caliente, más fría ésta. No existe mezcla entre los isótopos de uno y otra. Ello implica que el núcleo consume su hidrógeno y se convierte la estrella en gigante roja. Se sucederán nuevas reacciones nucleares a través de las cuales irán apareciendo nuevos elementos, hasta llegar al isótopo Fe-56. En sus últimas fases de evolución, ese cuerpo celeste arrojará los elementos creados en una explosión de supernova.

¿Qué procesos de nucleogénesis engendran los elementos más pesados que el hierro? Digamos primero que se forman por captura sucesiva de neutrones en las postrimerías de la evolución estelar o durante los episodios de explosión. Para ello, es imperioso que, en la fase de secuencia principal, el carbono, oxígeno y nitrógeno se conviertan en N-14. Este isótopo, en un momento precoz de la fase de gigante roja, atravesará una reacción de la que resultará O-18, portador de dos neutrones. Los neutrones se enlazarán con los elementos pesados, favorecidos por su sección eficaz, mayor que la de los elementos ligeros. Fowler sugería dos procesos, el *s* (de *s*low, lento), en virtud del cual los isótopos pesados que habían surgido por captura de un neutrón solían desintegrarse antes de incorporar otro, y el proceso *r* (de *r*apid, rápido), en que el flujo abundante de neutrones permitía la generación de núcleos más pesados que el bismuto. Sometió esta elegante trabazón teórica a ensayo experimental en el laboratorio de Oak Ridge.

Puestos a explicar la nucleogénesis de los elementos químicos, ¿por qué no empezar por el principio? En la gran explosión inicial que dio origen a nuestro universo, no pudo nacer ningún elemento más pesado que el helio. Surgieron en estrellas ya constituidas. Las más viejas de ellas se encuadran en lo que los astrofísicos llaman Población II, cuyos espectros delatan una inmensa pobreza en elementos más pesados que el helio. Debido a su avanzado estadio de evolución –la mayoría se encuentran en fase de gigantes rojas– poseen, además, menos opacidad.

Conocida la historia química de las estrellas, aventuró una cronología del cosmos. En particular, demostró que las proporciones de isótopos observadas en el caso del uranio y torio eran coherentes con la síntesis de estos materiales a lo largo de 7000 millones de años y no encajaban en un marco teórico de un proceso que fuera más precipitado. Y asignó a la galaxia una antigüedad de 20.000 millones de años, edad que otras pruebas han rebajado un poco. En otro orden, extiende sus cálculos más allá de las estrellas que centuplican la masa solar, y sugiere la existencia de otras que podrían multiplicar hasta por 10^8 la masa del Sol: esas estrellas supermasivas podrían constituir la energía creadora de las explosiones que se producen en los núcleos galácticos, origen de las intensas radiofuentes. A los pocos meses (1963) de esa hipótesis, se descubren los quasars, que hallan fundamento teórico en las estrellas supermasivas.

Esta teoría astrofísica que le ha valido el Nobel a Fowler tuvo su guión en un artículo titulado "Synthesis of elements in stars" (*Review of Modern Physics*, 29, 1957, pp. 547-650), firmado por Fred Hoyle, Geoffrey R. Burbidge y Margaret Burbidge, famoso ya y conocido por las iniciales BBFH. Los Burbidge aportaron, sobre todo, su labor observacional. Con Hoyle, Fowler ha trabajado en constante relación: años sabáticos, veranos y múltiples visitas mutuas para atar cabos y dar rienda suelta a esa imaginación que constituye el caballo de arrastre de la ciencia. (Sobre el olvido de Hoyle a la hora de conceder el Nobel a este nuevo campo de la astronomía, *Nature* se dolía en un reciente editorial de que el comité otorgador haya querido librarse de antemano de posibles críticas a propósito del origen extraterrestre de la vida en nuestro planeta, defendido con ardor por el astrónomo inglés.)

Para ponderar el temple de Subrahmanyan Chandrasekhar, los biógrafos se recrean en la humillación pública a la que le sometió Arthur Eddington, el más influyente de los astrónomos, a mediados de la década de los treinta, en la sesión de la Unión Mundial de Astrofísicos cuando rechazó por falso el trabajo sobre los agujeros negros. Chandrasekhar estaba entonces en Cambridge y no dudó dos veces cuando la Universidad de Chicago le ofreció en 1937 un puesto.

Nacido en Lahore (actual Pakistán) el 10 de octubre de 1910, estudió física en el Presidency College de Madrás. A los veinte años se trasladó a Inglaterra, para especializarse en astrofísica teórica con R. H. Fowler, quien había aplicado la mecánica cuántico-estadística (de Fermi-Dirac) a la materia densa de las estrellas enanas. Pero era éste un trabajo teórico que necesitaba sustentación matemática, muy compleja por otro lado. En su tesis doctoral de 1933, Chandrasekhar demostró que la densidad de la materia de esas estrellas debía variar con la presión en una razón determinada e independiente de la temperatura. (Adviértase que las estrellas

en cuestión amasan tanta materia como el Sol en un volumen un millón de veces menor, lo que explica su altísima densidad, del orden del millón o cien millones de gramos por centímetro cúbico.) He aquí, resumida, su argumentación: la razón de presión a la densidad de los electrones es del mismo orden que la energía cinética media por electrón. En un marco no cuántico, esta energía es térmica; en el de la estadística de Fermi-Dirac, la energía cinética es del orden de la energía de Fermi, de los electrones. Esta energía de Fermi es función de la densidad, pero lo que esa función signifique dependerá de la relación dinámica entre energía cinética y el momento. Si la masa estelar es suficientemente pequeña, la presión mecánica centrífuga y la gravitación centrípeta se equilibrarán; la evolución final de una tal estrella será la muerte tranquila, enfriándose lentamente hacia un estado terminal de temperatura cero. Pero si la masa de la estrella es notable, por relatividad especial se demuestra que la presión mecánico-cuántica no puede competir con la gravitación ni con la presión térmica; en el límite, la estrella seguirá contrayéndose a medida que radia energía y terminará por tornarse invisible. Si la contracción entra dentro del radio de Schwarzschild, por relatividad general se demuestra que no puede escapar ya radiación de la estrella: se forma un agujero negro.

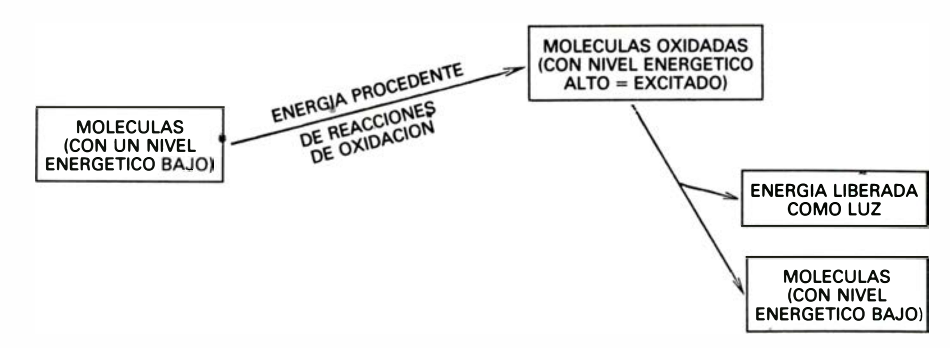
Chandrasekhar había obtenido en su primer año en Cambridge un resultado sorprendente: hay una masa máxima finita para cualquier enana blanca ideal. Masa que lleva su nombre y que se la supone de 1,4 masas solares para la estrella de hidrógeno y de 1,2 masas solares para las enanas estables de hierro. En el límite, contribuiría al desarrollo de la teoría de la nucleogénesis: refuerza la hipótesis de que, en las últimas fases de la evolución de una estrella masiva, el colapso del núcleo degenerado de la misma produciría el suficiente número de neutrones que se requerían en los procesos rápidos (*r*).

Suele cambiar periódicamente de campo de interés, si bien a modo de extensión de sus dos preocupaciones principales: estructura y evolución de las estrellas y agujeros negros; verbigracia, la teoría de las perturbaciones de estos últimos, que trata de la interacción entre los agujeros negros y su entorno, la dispersión y absorción de los distintos tipos de ondas incidentes. En 1952 se hizo cargo del *Astrophisical Journal*, elevándolo al prestigio y rigor de que hoy goza. Entregado a sus clases, aun cuando el número de alumnos no sea más que un par, como en aquel curso formado por T. D. Lee y Frank Yang, quienes recibirían conjuntamente el Nobel de física en 1957. De su última obra –*The Mathematical Theory of Black Holes*, 1983– ha escrito Roger Penrose en una sucinta reseña: "se trata de una obra maestra".

Quimioluminiscencia y bioluminiscencia

En el fenómeno de la fluorescencia, la fuente de energía que determina el "salto" de unas moléculas a un estado excitado es una radiación incidente; al descender de ese estado excitado dichas moléculas, aparece la emisión de la radiación fluorescente, de longitud de onda mayor que la que tenía la incidente.

Este fenómeno fue estudiado en disoluciones de sulfato de quinina por el matemático inglés Stokes. Por mostrar este compuesto reflejos a la luz, de tono azulado, similar a los que presenta el mineral fluorita, Stokes propuso, hacia mediados del siglo XIX, utilizar el nombre de fluorescencia para designar



1. Esquema general del funcionamiento del proceso quimio- o bioluminiscente

LUMINOL + H_2O_2 → ACIDO α-AMINOFTALICO + N_2 + LUZ

2. Esquema correspondiente a un ejemplo del proceso de quimio-luminiscencia

LUCIFERINA REDUCIDA + ATP + O_2 —LUCIFERASA→ OXILUCIFERINA + CO_2 + AMP + PP + LUZ

3. Esquema correspondiente a un ejemplo del proceso de bioluminiscencia

VALORACION DE ENZIMAS	VALORACION DE SUSTRATOS, COENZIMAS, METABOLITOS Y OTROS COMPUESTOS
ATPasas	ATP, ADP, PP, AMP, AMPc
CREATINQUINASA	CREATINA, CREATINFOSFATOS
PIRUVATOQUINASA	PIRUVATO, FOSFOENOLPIRUVATO
HEXOQUINASA	HEXOSAS (ALGUNAS)
GLUCOSAOXIDASA	GLUCOSA
PEROXIDASA	H_2O_2
AMINOACIDOXIDASAS	L-AMINOACIDOS
ALCOHOLDESHIDROGENASA	ETANOL
LACTATODESHIDROGENASA	NADH
ISOCITRATODESHIDROGENASA	NADH
GLUCOSA-6-FOSFATODESHIDROGENASA	NADH
	ACIDO LACTICO, TRIGLICERIDOS, UREA, VITAMINAS, HORMONAS, MEDICAMENTOS; RADICALES LIBRES

4. Ejemplos de enzimas y sustancias valorables por quimio- o bioluminiscencia. Ensayos basados en técnicas quimio- o bioluminiscentes

lo que antes se venía considerando como un fenómeno de "dispersión".

Si la emisión de energía tiene lugar de manera prolongada durante algún tiempo después de la absorción luminosa inicial, se habla de fosforescencia. Tanto la fluorescencia como la fosforescencia pueden incluirse dentro de la denominada luminiscencia; o más exactamente, de la fotoluminiscencia. Sin embargo, últimamente tiende a reservarse el uso de la palabra luminiscencia a los fenómenos de quimioluminiscencia y de bioluminiscencia. (El término "Lumineszenz" fue propuesto por el alemán Wiedemann, en 1880.)

En el caso de la quimioluminiscencia, la energía necesaria para pasar las moléculas al estado de excitación procede de una reacción química. Si en esta reacción participan agentes de tipo biológico (catalizadores como los enzimas, de procedencia habitualmente biológica), el fenómeno se denomina bioluminiscencia. En cualquiera de estas modalidades, el descenso energético de esas moléculas, momentáneamente excitadas, se acompaña de emisión de luz; por este motivo se le da el nombre de luminiscencia. (En otros casos también podrá liberarse esa energía en forma de calor.)

La observación de fenómenos bioluminiscentes está mencionada desde muy antiguo. El propio Aristóteles describe la existencia en algunos peces y hongos de lo que hoy llamamos bioluminiscencia. Pero hasta época reciente, en el siglo XIX, no se aborda el mecanismo del proceso. Los trabajos efectuados en 1880 con extractos del insecto "gusano de luz" por Dubois fueron decisivos. Actualmente se considera que, exceptuados los metales en estado sólido, numerosas sustancias son fluorescentes en mayor o menor grado, según las circunstancias.

Se ha podido deducir que existen diversos grupos de sustancias químicas (diacilhidrazidas como el luminol, sales de acridinio como la lucigenina, oxalatos de diarilo o la lofina y la siloxina), dotadas de estructuras muy distintas unas de otras (frecuentemente pertenecientes a la serie aromática), capaces de participar en los fenómenos de tipo quimioluminiscente. Lo mismo que existen grupos cromóforos en las sustancias colorantes, se ha intentado descubrir la presencia de grupos fluoróforos en las sustancias luminiscentes.

En los océanos, unos dos tercios de los organismos (fotobacterias) que viven en la capa de los 2000 metros más superficiales producen bioluminiscencia, aunque puede presentarse también este fenómeno en zonas profundas (en peces abisales), no conociéndose bien aún su cometido biológico. Son responsables los enzimas luciferasas, hallados en bacterias, hongos, insectos, gusanos, caracoles marinos, almejas y crustáceos.

La determinación de numerosas actividades enzimáticas acoplables, más o menos directamente, a sistemas como el del $NAD(P)^+ \rightleftharpoons NAD(P)H + H^+$ permite, además, la valoración de muy distintos compuestos o enzimas. (Obviamente, esta apreciación es válida también para otras técnicas tales como

las espectrofotométricas, donde se aprovechó inicialmente este hecho en el ensayo clásico y generalizado del NADH a 340 nanómetros; así como para las fluorimétricas.)

Un aspecto particularmente valioso de las metódicas bioluminiscentes es el de su altísima sensibilidad, similar (y a veces superior) al de las técnicas isotópicas, manteniendo sin embargo gran selectividad y especificidad.

Tres hechos han contribuido a restringir el uso de las metódicas bioluminiscentes: la necesidad de utilizar reactivos purísimos, la demora en la comercialización de los enzimas a emplear en estos ensayos (había que purificarlos antes de cada serie de experimentos) y la exigencia de aparatos especiales para asegurar la sensibilidad y exactitud propias de estas técnicas, aunque pueden usarse también para estas medidas (con algunas limitaciones) contadores de centelleo.

Vale la pena resaltar algunas aplicaciones de las técnicas luminiscentes dentro de la investigación bioquímica: fosforilación oxidativa y fotofosforilación (mediante determinación de ATP en la mitocondria y el cloroplasto, respectivamente), estudios de metabolismo intermedio (por la determinación de nucleótidos y algunos enzimas relacionados con estas reacciones) e investigación de procesos metabólicos asociados a membranas, núcleo y lisosomas celulares.

El empleo del método bioluminiscente ha abierto amplias perspectivas en distintos campos. Teniendo en cuenta, por ejemplo, que las bacterias contienen aproximadamente 2×10^{-16} g de ATP por célula, puede determinarse por bioluminiscencia el número de bacterias, no sólo en muestras de sangre o de orina, sino también en alimentos (leche, productos lácteos, helados y carnes) con mayor rapidez y sensibilidad que por los procedimientos bacteriológicos tradicionales. Más aún. Determinado el número de bacterias en suero sanguíneo por bioluminiscencia, puede deducirse el grado de eficacia de un tratamiento con antibióticos.

Cabe, asimismo, conectar la técnica luminiscente con otras. Piénsese en las de inmunoensayo (inmunoensayo quimioluminiscente), inmovilización de enzimas, automatización (para determinación de etanol) y autolumografía. Esta última técnica, que presenta cierto paralelismo en su fundamento y aplicaciones con la autorradiografía, consiste en la localización de una actividad enzimática mediante la detección de los fotones emitidos en la misma. (José A. Cabezas Fernández del Campo.)



Arroz

De la familia de las gramíneas, como el trigo y el maíz, en él funda su alimentación una extensa fracción de la población humana. Los avances en genética han aumentado notablemente la producción por hectárea

M. S. Swaminathan

Entre las principales plantas de las que se aprovecha el grano, el arroz es la única que se cultiva casi exclusivamente para alimentación humana. Aporta la mitad de la dieta de 1600 millones de personas, y en él se funda de un cuarto a la mitad de la dieta de 400 millones más. Como cabría esperar, su siembra ocupa una extensa porción de tierra cultivada: unos 145 millones de hectáreas, lo que constituye el 11 por ciento de la superficie cultivable mundial. En el año arrocero que finalizó el 31 de julio de 1982 la cosecha alcanzó los 410,9 millones de toneladas métricas. De no ser por una serie de notables avances genéticos que han llevado al cultivo variedades de alta producción y resistentes a enfermedades y a plagas de insectos, la cosecha hubiera sido muy inferior, y millones de personas habrían padecido una grave desnutrición.

El arroz es una planta herbácea anual que pertenece a la misma familia que la cebada, la avena, el centeno y el trigo, con los que comparte muchas características. Se encuadra dentro del género *Oryza*. Sus orígenes se remontan tanto en el tiempo que quizá nunca puedan trazarse con certeza. Mi colega T. T. Chang, del Instituto Internacional de Investigaciones Arroceras, de Filipinas, ubica el hábitat original en el viejo hipercontinente de Gondwana. Las especies de *Oryza* derivarían hacia diferentes hábitats geográficos cuando Gondwana se fragmentó y dio origen a Africa, la Antártida, Australia, Madagascar, América meridional y sur y sudeste asiáticos.

Los dos arroces cultivados, con gran número de variedades cada uno, son *O. sativa*, de Asia, y *O. glaberrima*, de Africa occidental. El género *Oryza* incluye también 20 especies silvestres, repartidas por Asia, Africa, Australia y América Central y del Sur. Las pruebas arqueológicas sugieren que la domesticación de *O. sativa* comenzó en Asia hace más de 7000 años, produciéndose la de *O. glaberrima* en Africa algo más tarde.

La diversidad de ambientes donde se cultiva arroz apenas si tiene parangón en el reino vegetal. Surgió en los trópicos cálidos y húmedos, donde las lluvias monzónicas y las inundaciones creaban un ambiente acuático durante parte del año, al menos. La dispersión natural y la selección practicada por el hombre han extendido su cosecha desde las riberas del río Amur (53 grados de latitud norte), en la frontera entre la Unión Soviética y China, hasta la Argentina central (40 grados de latitud sur). Vemos campos en climas fríos de las altas montañas de Nepal y la India y en los cálidos desiertos de Pakistán, Irán y Egipto. Es cosecha de secano en Asia, Africa e Iberoamérica. Prosperan arroces flotantes en aguas de tres metros de profundidad en Bangladesh, Burma, India oriental, Tailandia y Vietnam. Se adapta bien a diferentes condiciones y se comporta mejor que otros cultivos para grano en áreas con suelos desfavorablemente salinos, alcalinos y sulfoácidos. A partir de estas características, la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO) prevé (en *Land Resources for Populations of the Future*) su extensión a nuevos países en las próximas décadas.

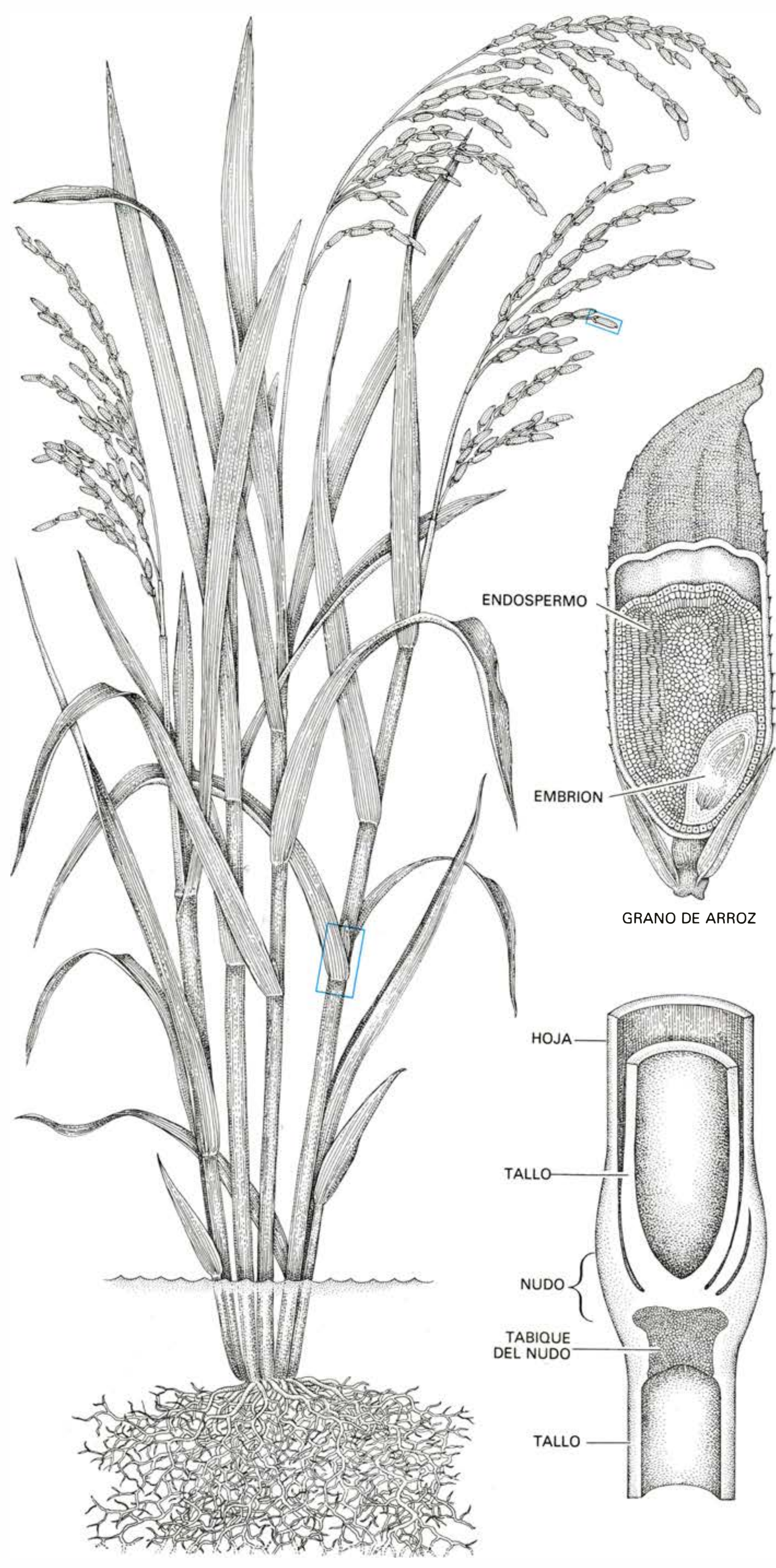
Los métodos de cultivo del arroz oscilan entre los sistemas muy mecanizados de occidente y aquellos que requieren abundante mano de obra, los de muchas zonas del sudeste asiático. La producción media va desde menos de una tonelada por hectárea en algunos países africanos hasta más de seis toneladas en Australia, Japón, Corea del Norte y del Sur y en los Estados Unidos. Durante siglos, la producción de Asia tropical fue de una a 1,5 toneladas por hectárea. Sin embargo, entre 1960 y 1980 las producciones medias aumentaron un 40 por ciento y la cosecha creció en más del 60 por ciento. En determinados países, así Indonesia y Filipinas, la cosecha y la productividad aumentaron más en esos 20 años que en los 7000 precedentes. Indonesia se autoabastece hoy casi por completo. Filipinas, hasta hace pocos años gran importador de arroz, exporta ahora frecuentemente por encima de las 100.000 toneladas anuales.

¿Cómo ocurrió esta transformación? El principal catalizador del cambio lo encontramos en el desarrollo de variedades de arroz de elevado rendimiento, apoyado en la tendencia hacia un tratamiento científico del suelo y de la sanidad de las plantas. Los gobiernos de varios países del sudeste de Asia implantaron políticas complementarias para acelerar el riego, la reforma agraria y aportación de fertilizantes, juntamente con modalidades de créditos liberales y precios subvencionados. El resultado fue un incremento en la producción que en la mayoría de los países superó el 5 por ciento anual.

Arroz asiático

A lo largo de milenios, la especie cultivada en Asia (*O. sativa*) se diferenció en tres subespecies instadas por las condiciones geográficas. Se trata de las subespecies sínica) y javánica (llamada también sínica) y javánica. Una clasificación adicional, que hace hincapié en el hábitat, en las peculiaridades de los suelos y disponibilidad de agua, es la de

1. HABILITACION DE TERRAZAS para plantaciones de arroz. Este sistema extiende enormemente el área disponible para su cultivo en las regiones montañosas de Asia. Las terrazas de la fotografía de la página opuesta corresponden a Filipinas. Se ofrecen cultivos en diferentes estadios de desarrollo, en arrozales con cinco a 10 centímetros de agua. La planta prospera en el agua gracias a un sistema de paso de aire que lleva oxígeno del tallo a la raíz. Estas terrazas con dique se inundan con agua de lluvia.



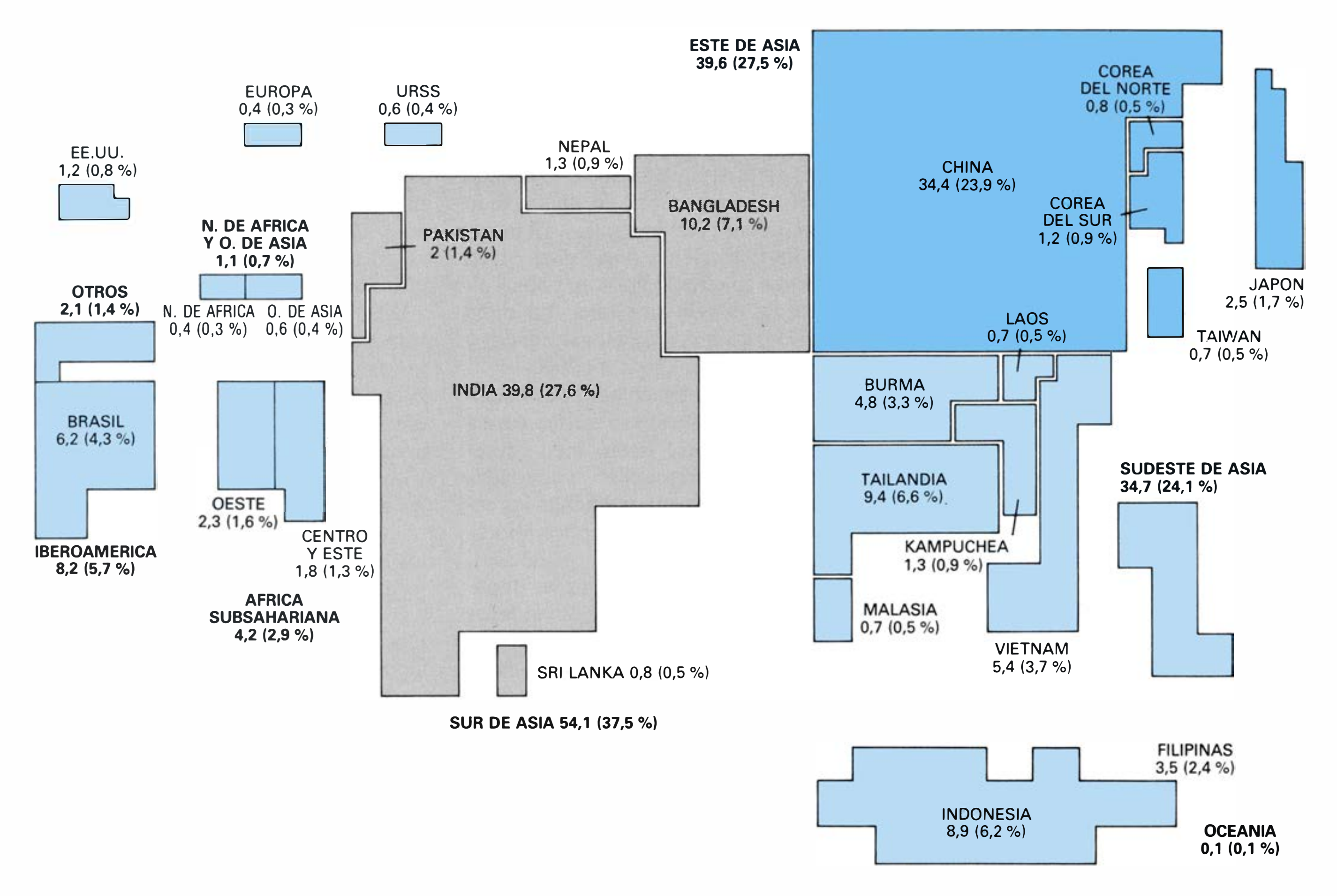
2. PLANTA DE ARROZ y secciones de un hijuelo, tallo y hoja, y de un grano. El arroz pertenece al género *Oryza*; esta planta es de la especie *O. sativa*, cultivada en Asia, de la que existen unas 120.000 variedades. La otra estirpe cultivada de arroz, de menor difusión, es la especie africana *O. glaberrima*.

arroces de secano de tierras altas y de tierras bajas y arroces de regadío y de aguas profundas. La subespecie índica estaba confinada originalmente a las regiones húmedas de los trópicos y subtrópicos asiáticos. La subespecie japónica se cultivaba en regiones de la zona subtropical templada. La javánica florecía en la región ecuatorial de Indonesia.

Pero no sólo difieren por su adaptación climática. Las tres razas se distinguen por ciertos rasgos del grano: contenido en amilosa (un derivado del almidón), elongación del mismo, temperatura a la que se vuelve gelatinoso y aroma al cocerlo. Las cualidades culinarias y gastronómicas del arroz vienen determinadas principalmente por las propiedades del almidón, que constituye el 90 por ciento del arroz descascarillado. La calidad del grano debe definirse también en relación con la localidad. Por ejemplo, para los tailandeses el arroz de buena calidad es el de granos largos, finos y traslúcidos que, a la cocción, dan un producto suelto y tierno. En el Japón se prefieren los de granos cortos y anchos, que en la cocción resultan pegajosos, lo que facilita su consumo con palillos. En algunas zonas de Pakistán y del noroeste de la India las variedades *basmati* de granos largos, finos y aromáticos alcanzan los precios más altos.

Siguiendo la expansión de la especie asiática cultivada, *O. sativa*, la mayoría de los investigadores están de acuerdo en que el área que ofrece la máxima diversidad de variedades es una faja que se extiende desde la región de Assam-Meghalaya, en la India, hasta las cordilleras del sudeste de Asia y del sudoeste de China. Las variedades de maduración temprana y que, por tanto, escapaban de los más graves efectos de la sequía periódica, emergieron probablemente hace de 15.000 a 10.000 años a lo largo de las laderas del sur y del norte del Himalaya. Las formas ancestrales anuales de *O. sativa* asomaron por la periferia de sus antecesores anuales silvestres, principalmente en las estribaciones meridionales del Himalaya y, en menor extensión, en el sur y sudoeste de China. La alternancia de períodos de sequía y la pronunciada variación de temperaturas aceleraron el desarrollo de las formas anuales de *O. sativa* en el nordeste y este de la India, el norte del sudeste asiático y el sur de China. Las primeras formas cultivadas eran todas índicas.

Las variedades cultivadas maduraban antes que sus antecesoras; sobrevi-



3. SUPERFICIE MUNDIAL DEDICADA AL CULTIVO DE ARROZ representada en un mapa esquemático. El tamaño de cada país pretende ser proporcional a la extensión (que se da en millones de hectáreas) del cultivo del arroz. También se da para cada país el porcentaje respecto al total mundial.

vían así mejor y producían más semillas en los climas más secos y fluctuantes. Tales factores favorecieron su expansión más al norte que las formas perennes. La creciente aridez del borde septentrional del Himalaya y sus cordilleras asociadas del sudeste asiático forzaron a los primeros habitantes de China a trasladarse hacia el este y el sur, hacia regiones más húmedas. Con la emigración humana, se dispersaron también las plantas, acelerando el proceso de diversificación ecogenética.

A partir de la raza primaria índica tropical se desarrolló una nueva raza para zonas templadas, dentro de los actuales límites de China, a lo largo de un brazo del río Brahmaputra (río Hsiangchuna): la raza japónica, así llamada desde que los trabajadores japoneses acuñaron el término en 1928. Chang sostiene que, puesto que los japoneses trajeron sus arroces de China, la japónica debería llamarse sínica (keng). Los arroces de talla alta y de grano verde grande emigraron a las islas de Indonesia desde la costa oriental del sur de Asia y se diferenciaron en la variedad javánica.

Los restos más antiguos de arroz cultivado, ambos índica, se encontraron en el este de China y el norte de la India; se remontan a 7000 años atrás. El arroz sínica más antiguo desenterrado en China data del 3300 a. de C. Granos de lo que probablemente era arroz silvestre se han encontrado en excavaciones de Non Nok Tha, en el norte de Tailandia, y se remontan a unos 4000 años a. de C. En Asia se tiene conciencia de la antigüedad del cultivo del arroz; a menudo se llama "agricultores de 50 siglos" a los arroceros.

La dispersión posterior de los cultivares asiáticos, principalmente índica, desde la India al Medio Oriente, norte de Africa y luego a Europa, comenzó probablemente unos 1000 años a. de C. También viajó el arroz desde la India a Madagascar y al este de Africa. Muchos cultivares javánica se abrieron camino desde Indonesia hacia Madagascar. Africa occidental obtuvo variedades de arroz de Europa o directamente del sur de Asia. Los países europeos proporcionaron la mayoría de los primeros cultivares sembrados en Sudamérica. Los Estados Unidos importaron semilla de arroz de Madagascar, sur de Asia y este de Asia.

No me he detenido en los arroces africanos cultivados (*O. glaberrima*) porque nunca han alcanzado la importancia económica del grupo asiático (*O. sativa*). R. Porteres, del Instituto de Investigaciones Agronómicas de París, postuló que la forma cultivada africana, *O. glaberrima*, se originó en el delta del río Níger. El centro primario de diversidad de *O. glaberrima* se extiende desde la zona pantanosa de la cuenca superior del Níger hasta dos áreas próximas a la costa guineana. Probablemente se cultivó arroz en dicha cuenca fluvial unos 1500 años a. de C. y en Guinea unos 400 años más tarde. En Africa occidental *O. glaberrima* constituye una cosecha predominante en áreas anegadas del Níger y del Sokoto.

La forma cultivada africana y su raza silvestre anual (*O. barthii*) divergen menos que sus contrapartidas asiáticas. Aunque Porteres reconoció en *O. glaberrima* dos subespecies, denominadas *vulgaris* y *humilis*, otros investigadores creen que el complejo *barthii-glaberrima* puede diferenciarse en dos ecotipos: de aguas profundas y de tierras altas. Los cultivares africanos son inferiores a los asiáticos en aguas pro-

fundas porque están limitados en su capacidad genética para elongarse en tales condiciones. La diferenciación de *O. sativa* en Asia precedió probablemente a la de *O. glaberrima* en Africa occidental, a juzgar por sus historias de domesticación y por la extensión de la diversificación varietal dentro de las especies.

Conservación de la herencia

La diferenciación ecogenética del arroz se realizó principalmente por selección natural y por selección artificial del hombre en diferentes ambientes, tras hibridación, segregación o mutación. Las múltiples razas de arroz así desarrolladas constituyen una valiosa reserva de germoplasma. Deben recogerse, conservarse y clasificarse para los programas de mejora actuales y futuros. Con esta finalidad, el Instituto Internacional de Investigaciones Arroceras ha fundado el Centro Internacional de Germoplasma de Arroz.

La totalidad del abanico de germoplasma de *Oryza* deriva de tres fuentes. En el primer grupo se encuadran las especies silvestres, los cultivares primitivos, los híbridos naturales entre variedades cultivadas y sus parientes silvestres. El segundo grupo incluye variedades comerciales, antiguas, de poco uso y especializadas. El tercero está formado por variedades que surgen por intervención humana, deliberada o no.

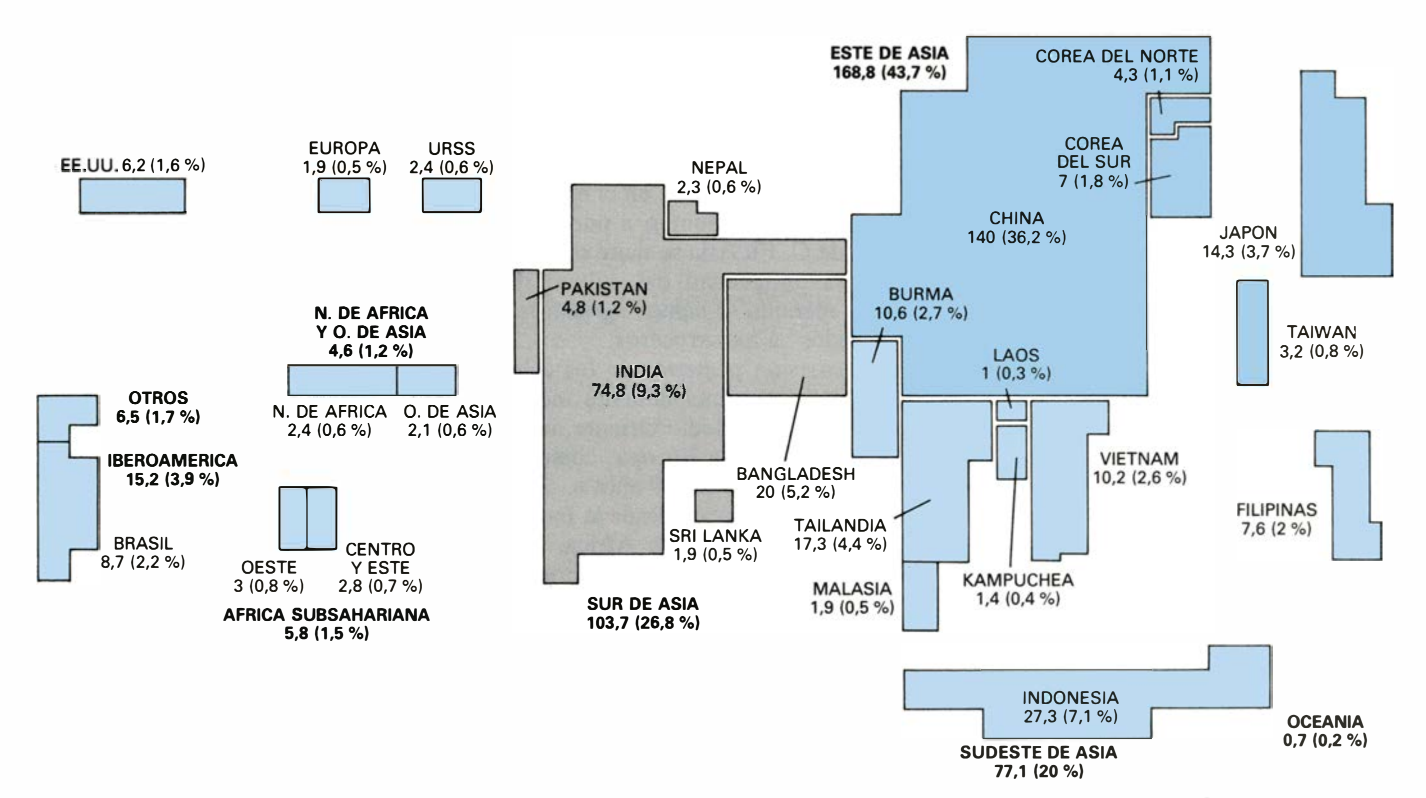
Se estima que el número de cultivares de arroz ronda la cifra de 120.000. Todos los países arroceros importantes de Asia reunieron, a partir de la década de 1930, germoplasma de un elevado número de especies indígenas. El tamaño de las colecciones nacionales oscila entre unos cuantos cientos en Laos y alrededor de 40.000 en China. La India ha catalogado unas 25.000 variedades y los Estados Unidos más de 7000.

La mayoría de las colecciones nacionales de Asia presentan varias características comunes. Están bien representadas las principales variedades comerciales, pero son pequeñas las colecciones de variedades de poca importancia y de variedades primitivas. Abundan en esas colecciones las duplicaciones, especialmente de variedades extranjeras que han conseguido reputación internacional. El número de estirpes de mejora que se conserva es pequeño y se han recogido pocos arroces silvestres. La falta de instalaciones refrigeradas para almacenar la semilla ha forzado a los centros nacionales a cultivar y renovar sus colecciones casi anualmente, lo que resulta en mezclas, pérdidas y errores de etiquetado, y en una carga de trabajo excesiva para las capacidades locales.

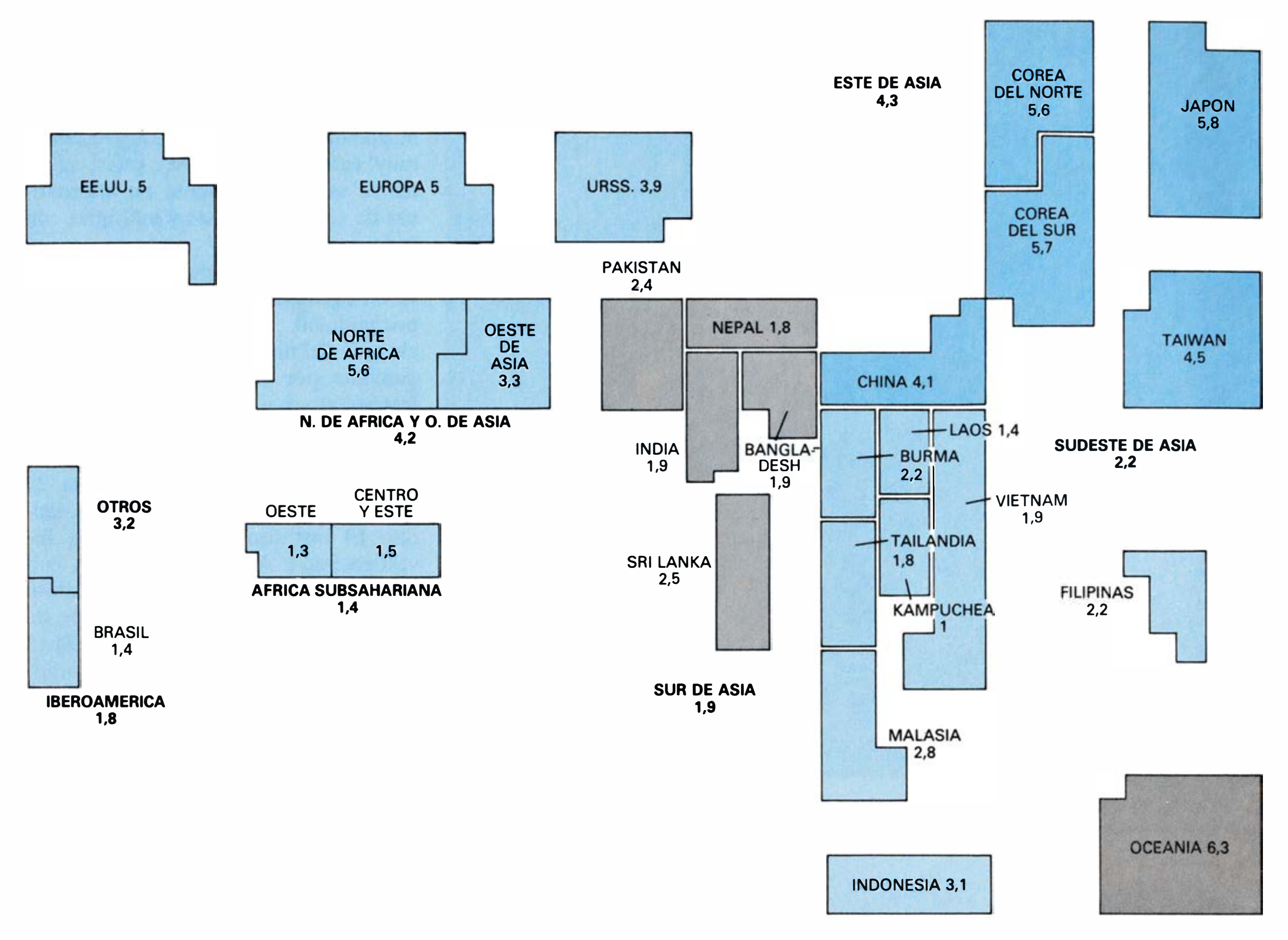
El Instituto Internacional de Investigaciones Arroceras estableció su centro de germoplasma en 1961. La colección alcanzó las 12.000 muestras en 1970, con la ayuda de una excelente cooperación por parte de varios centros nacionales. La instalación para almacenamiento en frío que se construyó al principio del programa ha convertido el banco en centro de intercambio de semillas entre países arroceros.

Desde 1971 el banco constituye el depósito central para la colección básica de los arroces mundiales. En 1983, la colección abarcaba 63.000 cultivares asiáticos, 2575 arroces africanos, 1100 arroces silvestres y 680 variedades que se conservaban para investigar caracteres genéticos. También se conservan miles de estirpes para mejora con uno o dos rasgos deseables.

Para proporcionar mayor seguridad a la colección, fruto de miles de años de selección natural y humana, el instituto deposita un juego duplicado de semillas en el Laboratorio Nacional de Almacenamiento de Semillas de Fort Collins, Colorado. Se acaba de poner en marcha un nuevo plan quinquenal, con la ayuda de investigadores arroceros nacionales y del Consejo Internacional de Recursos Genéticos Vegetales, para recoger y conservar las restantes 30.000 o 40.000 variedades. Cuando la cooperación de científicos, agricultores, estudiantes y personal militar con acceso a



4. PRODUCCION DE ARROZ, representada en millones de toneladas métricas. Se da el porcentaje de contribución de cada país al total mundial. Las producciones corresponden a arroz-cosecha, es decir, arroz trillado pero no descascarillado. Ha aumentado notablemente en lo que va de siglo.



5. PRODUCTIVIDAD ARROCERA, en toneladas métricas por hectárea cultivada. Este esquema y los dos precedentes se basan en trabajos de Adelita C. Palacpac, del Instituto Internacional de Investigaciones Arroceras, con datos del Departamento de Agricultura de los Estados Unidos y de la FAO.

áreas remotas termine la tarea, en 1987, el arroz será la primera de las plantas alimenticias de interés que se conserve para el siglo XXI y posteriores en una colección de variabilidad genética de ocurrencia natural lo más completa posible.

Un reciente ejemplo subraya la importancia de esta cooperación mundial. Hace dos años, el Instituto Internacional de Investigaciones Arroceras acusó recibo (a través de Oxfam, organización humanitaria con sede central en el Reino Unido) de una petición de Kampuchea: que se proporcionara una colección de variedades tradicionales de arroz que se habían cultivado anteriormente en aquel país. Muchas variedades locales se habían perdido durante la guerra. El banco de germoplasma del instituto tenía 800 variedades de arroz de Kampuchea, fruto de un plan de recogida que se inició en 1973. Se suministraron en 1981 y 1982 semillas de unas 140 variedades, de las que pedía el Ministerio de Agricultura, que ya se están multiplicando en ese país. De no ser por la colección de germoplasma del instituto se habrían perdido para siempre esas variedades tradicionales de adaptación local.

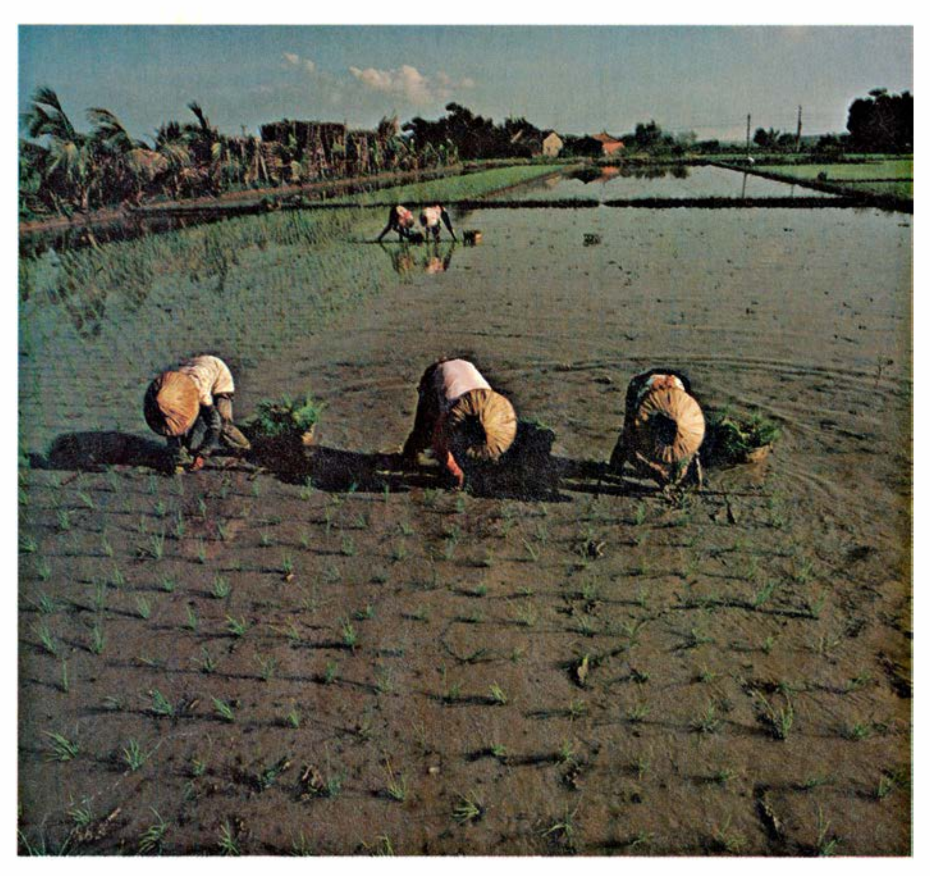
Otro ejemplo. Si los científicos y el Instituto Internacional de Investigaciones Arroceras no hubieran conservado algo de la variabilidad genética única del arroz cultivado en el nordeste del Himalaya y en el estado de Kerala, en el sur de la India, habría sido difícil adelantarse a la plaga de *Nilaparvata lugens*. Ese insecto constituye una amenaza para las variedades más productivas, al desarrollar rápidamente resistencia a los pesticidas y lograr vencer la resistencia genética de la planta de arroz, basada en genes simples. Los arroces de esta región poseen genes que confieren resistencia contra el insecto y también contra otras plagas.

Por qué se adapta el arroz

El arroz dispone de un eficaz sistema para el paso del aire desde el tallo hasta la raíz, lo que lo hace adaptable a una amplia gama de condiciones ambientales. El sistema permite su desarrollo en suelos anegados. El aire penetra en la planta a través de los estomas de los limbos y vainas foliares y pasa a los nudos de la base de la planta. A medida que el aire avanza del tallo a la raíz va suministrando oxígeno a los tejidos, donde se utiliza para la respiración. Desde las raíces el aire se difunde al suelo que las rodea. Para que el sistema funcione con eficacia al menos parte del tallo debe hallarse expuesto al aire.

Gracias a ese sistema de paso de aire, las raíces respiran aeróbicamente y aprovechan los carbohidratos para producir la energía que necesitan, aun cuando se cultiven en un ambiente anegado, anaeróbico. La mayoría de las plantas no pueden cultivarse en suelos anegados porque su sistema de paso de aire es menos eficaz. En el arroz, la eficacia del transporte de oxígeno del tallo a la raíz es 10 veces mayor que en la cebada y cuatro veces mayor que en el maíz.

En gran parte del mundo el arroz se



6. **SIEMBRA MANUAL en un arrozal de Taiwán. La escena es típica de los primitivos métodos de cultivo que aún se emplean en gran parte de Asia. En algunas zonas se echa la semilla desde aeroplanos.**



7. **COSECHA MANUAL de arroz cerca de Chaing Rai, en el norte de Tailandia. Escena que se repite en muchísimas regiones donde se cultiva este grano. Sin embargo, en California, donde se ha prestado gran atención a la mecanización de las faenas arroceras, se registra una tendencia hacia el empleo de máquinas cosechadoras; suelen disponer éstas de cadenas que les permiten operar en arrozales fangosos.**

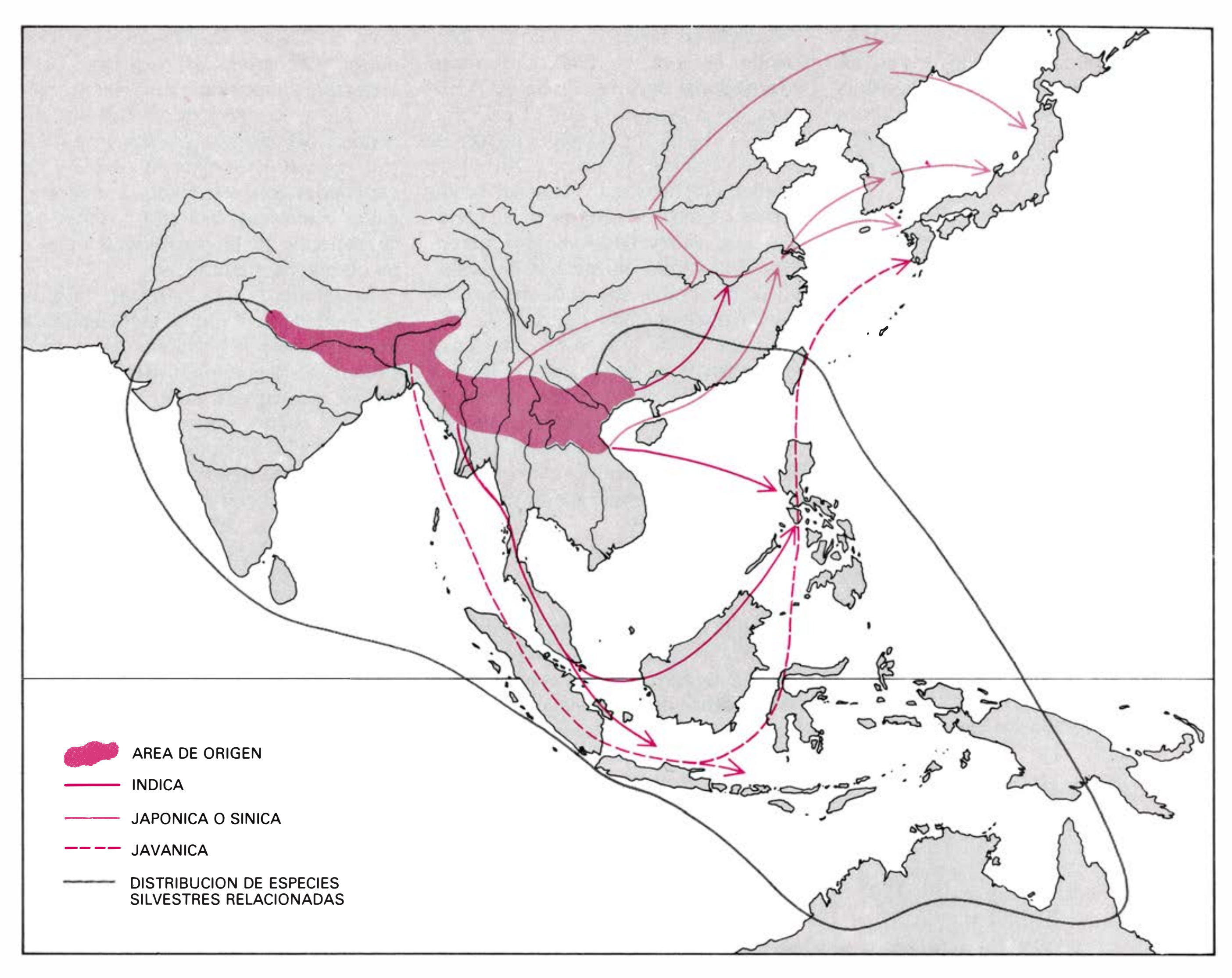
cultiva en suelos con una capa de agua de cinco a 10 centímetros. Pero la planta medra también en suelos secos de las tierras altas, donde se cultiva a veces lo mismo que el trigo y el maíz, y en zonas muy sumergidas. Puede, pues, plantearse satisfactoriamente en alternativas de cosechas mixtas y múltiples con trigo, maíz y soja.

El arroz flotante puede desarrollarse en aguas de 1,5 a cinco metros de profundidad. Resiste una elevación abrupta del nivel del agua y crece rápidamente por encima de él, llegando frecuentemente a alcanzar una altura de hasta seis metros. Se alarga normalmente a razón de dos a 10 centímetros por día, pero puede crecer hasta 25 centímetros por día en aguas profundas. El Instituto Internacional de Investigaciones Arroceras colabora con científicos tailandeses, de Bangladesh e indios para mejorar variedades de arroz adaptadas a áreas anegadas hasta un metro de profundidad, incorporando los genes para elongación del arroz flotante a variedades mejoradas semienanas de elevada productividad. Estos arroces de aguas profundas muestran una talla más baja en aguas someras, pero la desarrollan si sube el nivel de las aguas por inundación.

Fijación del nitrógeno

Aunque el 80 por ciento de la atmósfera sea nitrógeno, éste no lo utiliza directamente la planta de arroz. Se le hace accesible a través de la fijación biológica: conversión del nitrógeno gaseoso (N_2) en ion amonio (NH_4^+) que efectúan microorganismos especializados. En los suelos arroceros anegados se encuentran esquemas únicos de metabolismo microbiano, que incluyen transformaciones de nitrógeno y azufre, fijación biológica del nitrógeno y una rápida descomposición de pesticidas. La fijación biológica del nitrógeno es la más estudiada de esas actividades microbianas. Los suelos arroceros inundados estimulan a varios agentes fijadores: algas verde-azuladas de vida libre, bacterias fijadoras de nitrógeno y una relación simbiótica entre algas verde-azuladas fijadoras y el helecho acuático azolla.

La relación simbiótica fijadora de nitrógeno entre azolla y el alga verde-azulada *Anabaena azolla* alcanza su máxima actividad en suelos anegados, donde puede fijar hasta tres kilogramos de nitrógeno atmosférico por hectárea y por día. El nitrógeno acumulado por efecto de esta simbiosis se libera al



8. DIFUSION DEL ARROZ ASIATICO (*O. sativa*) desde su área de origen donde medró como herbácea silvestre anual. Las tres razas geográficas importantes de *O. sativa* son la índica, la japónica (o sínica) y la javánica. Los arroces índica se cultivaron originariamente en las zonas húmedas de los trópicos y subtrópicos asiáticos; las japónica en regiones templadas subtropicales y las javánica en la Indonesia ecuatorial. Se cree que el cultivo del arroz comenzó en Asia hace unos 7000 años. Este mapa se basa en los trabajos de T. T. Chang, del Instituto Internacional de Investigaciones Arroceras.

suelo y queda a disposición de las plantas de arroz cuando se descompone la azolla. Por su acusada actividad fijadora de nitrógeno la azolla ha servido de "abono verde" del arroz durante siglos en China y en Vietnam. (En este último país se ha dedicado un templo a la azolla.) El Instituto Internacional de Investigaciones Arroceras está llevando a cabo campañas conjuntas de investigación y formación para divulgar el conocimiento de la azolla y estimular su utilización como fuente de nitrógeno para el cultivo del arroz en todo el sur y sudeste de Asia.

En arrozales alcalinos y neutros medran espontáneamente algas verdeazuladas libres, fijadoras de nitrógeno, que forman en ellos una enorme biomasa. Hay en marcha diversos ensayos para mejorar su desarrollo a través de cultivos diseñados o por inoculación en los arrozales.

La fijación bacteriana es una tercera fuente de nitrógeno biológico para la planta de arroz; en los suelos de los arrozales hay bacterias fijadoras aerobias y anaerobias. Bacterias fijadoras de nitrógeno que se nutren del material vegetal muerto y de la materia orgánica liberada por las plantas vivas colonizan la superficie y las partes interiores de la raíz, así como las partes basales de los tallos. Por cada gramo de raíz de arroz (en peso seco) se cuentan unas 10^{10} células de bacterias fijadoras, que probablemente aportan una cantidad significativa de nitrógeno a la cosecha de arroz.

La planta necesita unos 20 kilogramos de nitrógeno por hectárea por cada tonelada de arroz producida. Las diferentes formas de fijación biológica de nitrógeno permiten al arroz rendir una o dos toneladas de grano por hectárea sin fertilización mineral suplementaria. Esta es una de las razones por las que los agricultores asiáticos han recogido durante siglos de una a dos toneladas por hectárea sin la aplicación de fertilizantes.

Rendir más

Las producciones de arroz quedaron estancadas hasta hace 20 años en la mayoría de los países asiáticos. Sólo en el Japón la cosecha mejoró gradualmente desde 1,3 toneladas por hectárea, en el año 900 de nuestra era, hasta 2,5 toneladas por hectárea a finales del siglo XIX, cuando la restauración Meiji impulsó la modernización de la agricultura. Las mejoras genéticas y prácticas agrícolas avanzadas aumentaron rápidamente las cosechas en el siglo XX. La producción actual de Japón, Corea del Norte, Corea del Sur, Australia y Estados Unidos (especialmente en Califor-

nia) da una media de unas seis toneladas por hectárea. En la mayoría de los países ha aumentado gradualmente el rendimiento a lo largo de los 15 últimos años, pero todavía hay muchos agricultores que sólo recogen dos toneladas por hectárea.

La razón genética de estos avances hay que buscarla en el desarrollo de variedades de arroz índica de elevada productividad. Los índica tradicionales eran de talla elevada con tendencia al encamado, es decir, a tumbarse por efecto del viento, la lluvia o el peso del grano. La mayoría de las variedades modernas son semienanas.

Los arroces índica semienanos modernos se obtuvieron a partir de mutantes semienanos, como la variedad china Dee-geo-woo-gen. A principios de los años cincuenta se desarrolló y cultivó en Taiwán la variedad Taichung Native 1, el primer arroz índica semienano obtenido fuera de la China continental. Era resistente al encamado, respondía bien a las aplicaciones de fertilizantes y se mostraba bastante insensible a las diferencias de insolación con la latitud.

Cuando las fundaciones Ford y Rockefeller crearon, en 1960, el Instituto Internacional de Investigaciones Arroceras, en cooperación con el gobierno filipino, uno de los primeros proyectos que se pusieron en marcha fue el de obtener drásticos aumentos del rendimiento a través de programas de mejora con utilización de los mejores parentales disponibles en muchas naciones. Un éxito notable fue el desarrollo del arroz semienano IR8. Lanzada en Filipinas en 1966, esta nueva variedad batió marcas de producción y llegó a conocerse por el "arroz del milagro". Le siguieron otros semienanos de elevada producción: IR5, IR20, IR22 e IR24. Se caracterizan por ciclos progresivamente más cortos y mejor respuesta a buenas prácticas.

Insectos y enfermedades

La difusión de estas variedades de ciclo corto y elevada productividad animó a los agricultores a cultivar dos o más cosechas de arroz anuales en regadío y con elevadas aplicaciones de fertilizantes. Las nuevas técnicas proporcionaron más grano al mercado, pero arrastraron también mayores problemas fito y zoopatológicos. Los investigadores del instituto comenzaron a incluir en sus estrategias la creación de variedades con resistencia a enfermedades e insectos; hicieron hincapié en la reducción de la variabilidad anual y en un mayor rendimiento.

Resultado de esta estrategia ha sido la variedad IR36, que se cultiva en más de 10 millones de hectáreas de los arrozales de todo el mundo. Resiste cuatro de las principales enfermedades del arroz y cuatro importantes insectos, entre ellos los biotipos 1 y 2 de *Nilaparvata lugens*. Se desarrolla bien en diversos ambientes culturales, tolera condiciones adversas del suelo, tiene grano de buena calidad y madura en 110 días, lo que permite recoger hasta tres cosechas al año en los arrozales de regadío. IR36 desciende de 13 variedades procedentes de seis diferentes naciones. Entre sus parentales se cuentan IR8, el Taichung Native 1 y una especie silvestre de la India, *O. nivara*.

La mejora de la resistencia a plagas se complica por la adaptación de los insectos y organismos patógenos a las variedades de arroz reforzadas. Se han identificado así tres biotipos diferentes de *Nilaparvata lugens* que atacan directamente a la planta y que son también vectores de enfermedades víricas importantes. Cuando empezó a cultivarse IR8 sólo se conocía un *Nilaparvata lugens* que atacaba al arroz. Desde entonces, a medida que se obtenían nuevas variedades reforzadas, se han multiplicado selectivamente biotipos resistentes del insecto. El biotipo 3 de *Nilaparvata lugens* puede haber comenzado ya a atacar arroces de ciertas áreas de Filipinas y de Indonesia. Afortunadamente, la variedad IR56 resiste al biotipo 3 y se ha extendido en esas áreas.

Es esencial que la investigación vaya por delante en la interacción dinámica entre las variedades de arroz y la resistencia de las plagas. Con esta finalidad, los investigadores del programa multidisciplinario para evaluación genética y utilización del instituto criban sistemáticamente las colecciones mundiales de germoplasma para definir su reacción frente a las principales plagas del arroz.

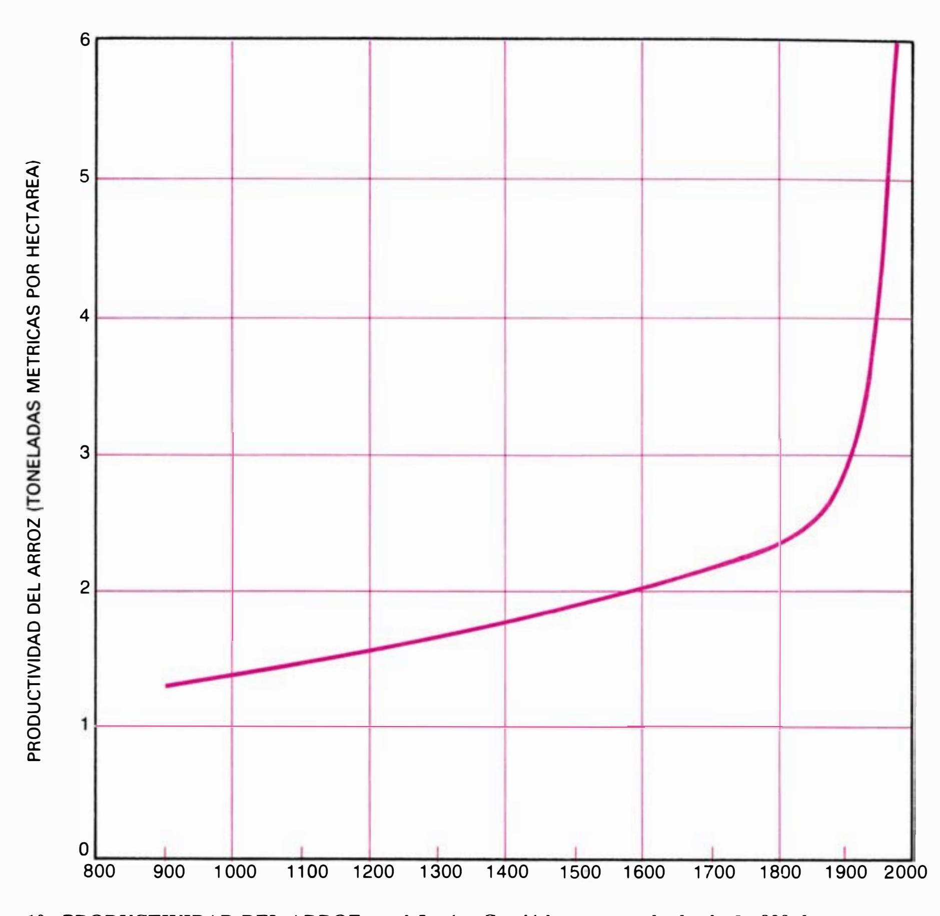
Las variedades con adecuada resistencia se cruzan en los programas de hibridación al objeto de combinar en una variedad muchas características favorables. Cada año el instituto realiza por encima de los 4000 cruzamientos. El material más prometedor se ensaya en

9. MERCED A LA FIJACION DEL NITROGENO por microorganismos del suelo puede obtenerse, sin aplicación de fertilizantes minerales, una cosecha de arroz de unas dos toneladas por hectárea. Entre esos organismos se cuentan el alga verde-azulada de vida libre *Anabaena azolla*, a la misma alga en asociación con el helecho acuático azolla y diversas bacterias. Se representan también sus hábitats ecológicos.

más de 300 localidades a través del Programa Internacional de Ensayos de Arroz, en el que intervienen unos 800 científicos de 70 países. A través de esta red cooperativa se han obtenido algunas variedades de interés.

Se vigila constantemente la aparición de nuevos problemas fito y zoopatológicos, lo que ha permitido identificar una nueva estirpe del virus de achaparramiento, bautizada provisionalmente con el nombre de RGSV2. Esta estirpe, que se encuentra en diversas zonas de Filipinas, puede causar graves daños a la planta. Se está realizando un intenso cribado para identificar fuentes de resistencia a la misma.

La lucha sin fin entre plagas y seres humanos ha llevado a la conclusión de que la meta de la agricultura debería ser el control de las plagas, y no su eliminación. Para controlar las plagas sin que se produzcan graves pérdidas de cosecha es necesario adoptar una estrategia integrada que incluya métodos de control genéticos, agronómicos y biológicos, juntamente con un prudente despliegue de pesticidas químicos inocuos y eficaces.



10. PRODUCTIVIDAD DEL ARROZ en el Japón. Creció lentamente desde el año 900 de nuestra era hasta finales del siglo XIX, y mucho más rápidamente a partir de entonces. El crecimiento gradual del primer período se debió a la generalización del riego y a las medidas de control de la inundación. La drástica subida reciente se explica por la obtención de variedades mejoradas (de mayor productividad y resistencia a plagas), la tendencia a la aplicación de fertilizantes minerales y una mejor práctica agrícola.

Perspectivas

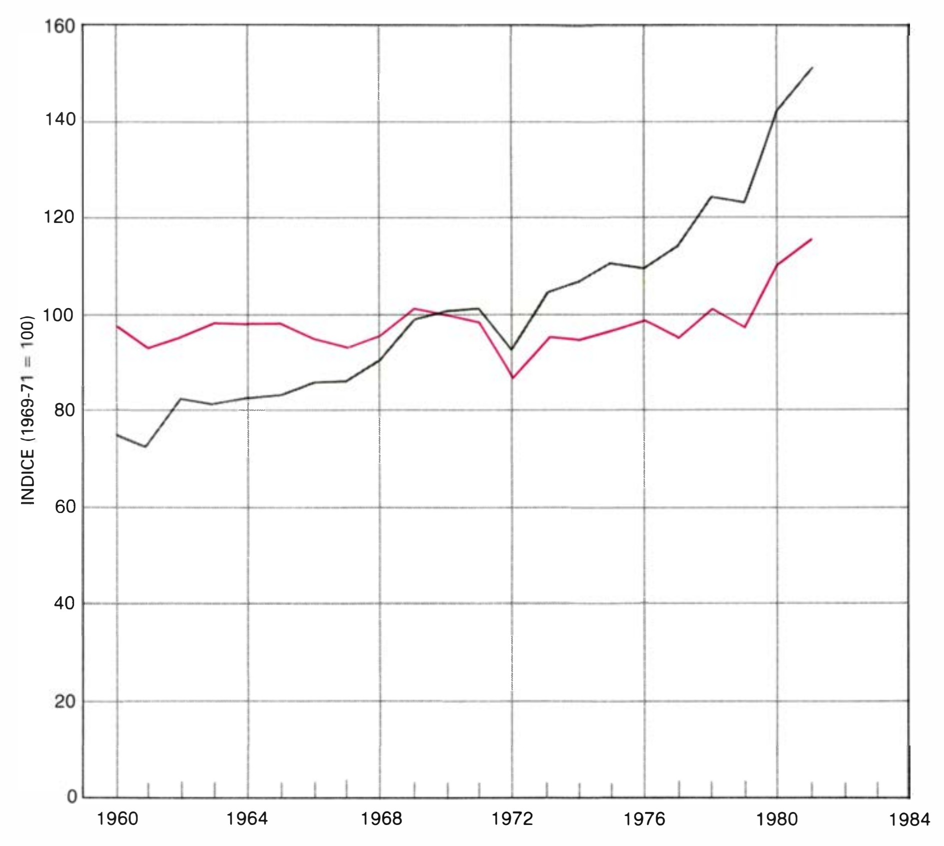
La Organización para la Alimentación y la Agricultura ha estimado que para contrarrestar el aumento de población será necesario un incremento anual del tres por ciento en la producción de arroz durante los años que quedan del siglo XX. Seguramente pueda incrementarse la producción a ese ritmo mediante la transferencia de técnicas ya disponibles, porque la tecnología moderna y los avances agronómicos no se han diseminado todavía plenamente en el mundo arrocero. Sería interesante, por tanto, conocer cuánto arroz puede producirse explotando el elevado nivel técnico de que se dispone actualmente. Esta información nos definiría los objetivos de las futuras producciones. En estaciones experimentales japonesas se han conseguido producciones de 9,3 a 10,2 toneladas por hectárea con la aplicación de técnicas agrarias modernas, pero incluso esas producciones son inferiores al máximo alcanzado: en 1960, un experto agricultor japonés obtuvo 13,2 toneladas por hectárea.

Para desarrollar planes nacionales realistas de producción de arroz será útil saber cuál es la producción potencial teórica de cada país. Dado que las plantas verdes convierten la energía solar en energía química, merecería la pena estimar la productividad potencial del arroz a diferentes niveles de la radiación solar incidente. Concepto importante, útil para la estimación de la productividad potencial del arroz, es el de la eficacia en la utilización de la energía, que puede asimilarse a la eficacia en la conversión de la energía por las células solares de silicio.

El arroz sintetiza la mayor parte del almidón del grano en el último estadio de su desarrollo, conocido por período de llenado del grano. Para estimar la productividad potencial del arroz en Asia tropical se supone que la eficacia de la utilización de la energía en el período de llenado del grano es del 3 por ciento para la cosecha de la estación húmeda (en la que se cultiva la mayor parte del arroz) y del 2,5 por ciento para la estación seca (cosecha fuera de estación, de cultivo con riego). En muchas localidades de Asia tropical es corriente una radiación solar de 300 calorías por centímetro cuadrado y por día durante el período de llenado del grano en la estación húmeda. En la estación seca, con menos días nublados, podrían producirse 12,5 toneladas por hectárea con una eficacia en la utilización de la energía del 2,5 por ciento, bajo una radiación solar incidente de 550 calorías por centímetro cuadrado y por día.

Otra cuestión de interés general: el potencial relativo de productividad del arroz en regiones de la zona templada y de los trópicos. El número de días para llenado del grano del arroz está determinado principalmente por la temperatura. Tarda unos 25 días en los trópicos y 35 días en la zona templada. Si la radiación solar incidente diaria es la misma para ambas regiones, la cantidad total de radiación solar incidente sería proporcional al número efectivo de días requeridos para el llenado del grano. En otras palabras, la productividad potencial del arroz sería un 40 por ciento más alta en las regiones de la zona templada que en los trópicos, dado que la radiación solar incidente es la misma. Por otra parte, pueden sembrarse dos o más cosechas de arroz en los trópicos si se dispone de agua suficiente.

Los cultivos de arroz en regadío suelen ocupar alrededor de un 30 por ciento del área de producción en los países



11. TENDENCIAS RECIENTES en la producción total de arroz (*gris*) y en la producción *per cápita* (*color*) en el sudeste de Asia. A pesar del continuo incremento en la producción total, y debido a la velocidad de aumento de la población, la cantidad de arroz *per cápita* no ha variado mucho desde 1969.

tropicales en vías de desarrollo y rinden un 50 por ciento del arroz obtenido en esos países. El arroz de regadío adquirirá importancia creciente en el futuro, pues la mayoría de los países en vías de desarrollo ponen cada vez más tierras bajo riego asegurado. El riego no sólo permite mayor productividad, sino también una mayor intensificación del cultivo, que se consigue sembrando en un mismo año una cosecha de arroz con riego, una cosecha de leguminosa de grano o alguna otra planta, sin riego, y una segunda cosecha de arroz.

En los campos arroceros de secano de llanura y montaña, las producciones son menores e inestables, puesto que la cantidad y distribución de las lluvias son irregulares y erráticas. La imprevisión de las condiciones climáticas hace arriesgado el cultivo de arroz en secano, de ahí que los agricultores inviertan poco en abonos, pesticidas y otros complementos de su compra. La producción media del arroz de secano se sitúa entre 0,5 y 1,5 toneladas por hectárea en las tierras altas; es inferior a dos toneladas por hectárea para el arroz de secano de tierras bajas. La producción ha crecido poco en las últimas dos décadas, especialmente en las tierras altas. El Instituto Internacional de Investigaciones Arroceras y otras instituciones están organizando intensos programas para aumentar y estabilizar la producción en las tierras sin riego.

Ingeniería genética

Los recientes avances en biotecnología quizás abran nuevos caminos para mejorar la productividad. Así, el cultivo de tejidos para obtención de plantas haploides: plantas dotadas de sólo un juego de genes, en lugar de los dos usuales. En estas plantas pueden estabilizarse características vitales, pues se elude la competencia entre dos genes de un mismo carácter. Además, esta técnica acorta, a tres generaciones en algunos casos, el tiempo necesario para obtención de una nueva variedad, que se cifra en 10 generaciones con los métodos convencionales de mejora. Los chinos han empleado este procedimiento para crear nuevas variedades de arroz, trigo y maíz.

La mayoría de las investigaciones sobre la técnica de los haploides se han llevado a cabo con arroz japónica en el Japón, China y los Estados Unidos. Hasta el momento no ha tenido éxito el procedimiento con el arroz índica, pues sus variedades no producen un número útil de callos, estados precoces de crecimiento en los que los tejidos no están diferenciados; este tejido es necesario para los cultivos. Se persigue con interés que la técnica sea aplicable al arroz índica.

Otra técnica prometedora es la de modificar las plantas induciendo mutaciones en los cultivos de tejidos. Células vegetales o granos de polen cultivados sirven como material mutable que produzca luego plantas completas. Una ventaja de este método sobre la inducción de mutaciones en semillas o en plantas adultas mediante irradiación o productos químicos estriba en que puede seleccionarse a nivel celular una característica determinada. Mediante esta técnica se han obtenido resultados prometedores para aumentar la tolerancia de la planta a la salinidad.

Entre los muchos objetivos de la ingeniería genética aparece como más ambicioso la incorporación en el arroz de genes para la fijación del nitrógeno. Sin embargo, el sistema de fijación involucra, al menos, 17 genes distintos, y no se han evaluado todavía las posibilidades de manipular tamaño número.

China ha introducido el arroz híbrido en el cultivo, abriendo la puerta a mejoras tan notables como las alcanzadas en el maíz a través de este procedimiento. No debe confundirse esta técnica con la de los cruzamientos, que consiste en combinar diferentes variedades para obtener una nueva línea estable. En la técnica de los híbridos solamente se siembra la primera generación (F_1), más vigorosa que sus dos parentales. El agricultor debe comprar nueva semilla F_1 cada año.

Resulta muy difícil aplicar la técnica con plantas autógamas, como el arroz, pero el reciente desarrollo de un sistema de mejora llamado de androesterilidad citoplásmica ha posibilitado las variedades híbridas en las autógamas. El sistema depende de la interacción de un gen que produce esterilidad en las plantas masculinas y de un factor citoplasmático de la célula sexual femenina. Se logró la aplicación de la técnica al arroz cuando investigadores chinos descubrieron plantas androestériles que podrían servir de parentales femeninos para cruzamiento con plantas masculinas adecuadas. La ventaja del híbrido, en términos de rendimiento, debe ser suficiente para compensar el elevado coste de producción de la semilla. No obstante, es demasiado pronto para predecir si tendrá o no éxito la técnica en China, o para valorar su aplicación en otros países.

Lagartijas unisexuales

Ciertas poblaciones de lagartijas de cola de látigo del continente americano están formadas únicamente por hembras. Este caso de reproducción por partenogénesis es insólito entre los vertebrados

Charles J. Cole

En 1958, Ilya Darevsky, zoólogo de la Academia de Ciencias de Leningrado, publicó un sorprendente informe sobre determinadas lagartijas de la Armenia soviética. Había encontrado varias poblaciones del género de saurios *Lacerta* formadas exclusivamente por hembras; y sugirió la posibilidad de que las lagartijas pusieran huevos viables, aunque no estuvieran fecundados. Muchos zoólogos recibieron con incredulidad el descubrimiento de Darevsky. Por aquella época la tesis según la cual toda especie de vertebrado se reproducía a través de la unión de los espermatozoides del macho con los óvulos de la hembra constituía punto menos que un axioma.

De un modo distinto reaccionaron algunos biólogos de los Estados Unidos ante el informe de Darevsky. Les resultaba verosímil y grato, a un tiempo. Sherman A. Minton, del Hospital Clínico de la Universidad de Indiana, y Richard G. Zweifel, del Museo Americano de Historia Natural, habían encontrado poblaciones de lagartijas de cola de látigo (del género *Cnemidophorus*), en los Estados Unidos sudoccidentales y México septentrional, que carecían de machos, aunque las poblaciones de *Cnemidophorus* vecinas estaban compuestas de machos y hembras en la proporción normal de uno a uno. Minton, Zweifel y otros biólogos más decidieron que ulteriores estudios de varias lagartijas de cola de látigo bien podrían añadir algún grupo a la lista de especies partenogenéticas.

Partenogénesis, o reproducción virginal, es un término que con frecuencia se aplica erróneamente a lo que sólo son formas aparentes de reproducción unisexual. Por ejemplo, en algunos peces unisexuales las hembras producen óvulos diploides: células germinales que contienen dos dotaciones completas de cromosomas. Los óvulos están completos desde el punto de vista genético, pero no se desarrollarán hasta que sean activados (no fecundados) por contacto con el espermatozoide de un macho. Puesto que no hay fecundación, se ha clasificado vagamente el proceso como si fuera de partenogénesis. Es más correcto denominarlo ginogénesis, término que destaca el papel dominante de la hembra. En otros peces unisexuales las hembras producen óvulos haploides: células germinales que, como las de las especies que se reproducen sexualmente, tienen sólo una dotación cromosómica. Tales óvulos no se desarrollan hasta que son fecundados por un macho perteneciente a una especie estrechamente emparentada, proceso que puede clasificarse de hibridogénesis para destacar el hecho de que los descendientes, todos hembras, son híbridos. Además, por lo menos una especie de pez, el rívulo (*Rivulus marmoratus*), se reproduce como un hermafrodita autofecundante. Sin embargo, en los tres casos la reproducción depende de la presencia de espermatozoides. Sólo la reproducción en ausencia segura de espermatozoides puede llamarse estrictamente partenogénesis.

Demostrar que no intervienen espermatozoides en la reproducción de una población libre es, evidentemente, imposible. No es fácil ni siquiera cuando se la confina en el laboratorio. Por ejemplo, tras el apareamiento, las hembras de alguna especie de reptil normalmente bisexual pueden almacenar esperma en su oviducto durante meses, o incluso años, y producir posteriormente huevos fecundados viables. Por tanto, para demostrar que algunas lagartijas de cola de látigo eran partenogenéticas estrictas era necesario criar hembras de la primera generación en ausencia de machos hasta que alcanzaran la madurez sexual, y luego ver si podían producir una segunda generación en ausencia continuada de machos.

A principios de la década de 1960 se capturaron lagartijas de cola de látigo aparentemente unisexuales y se criaron en laboratorio con el propósito de estudiar su reproducción. Los primeros intentos fueron descorazonadores: nadie sabía cómo mantener adecuadamente esos animales en cautividad. Algunas hembras pusieron huevos en el laboratorio, pero casi todos se malograron. Cuando se solucionó el problema en el laboratorio de Zweifel, nacieron *Cnemidophorus* de primera generación allí y en otros laboratorios. Surgió un segundo problema. Alrededor del 90 por ciento de las lagartijas acabadas de eclosionar moría en los tres primeros meses, y la mayor parte del resto en los seis primeros meses. Las lagartijas de cola de látigo precisan por lo menos un año para alcanzar la madurez sexual, de modo que empezaba a parecer imposible la obtención de *Cnemidophorus* de segunda generación.

Sin embargo, en 1973 se caracterizó un factor vital para la supervivencia de las lagartijas en cautividad. A los animales ya se les había proporcionado una superficie habitable, refugio, calor, alimento, agua y condiciones higiénicas apropiados. Sometidas a la iluminación normal del laboratorio, les faltaba la exposición a la radiación ultravioleta de la luz solar. La acción de las longitudes de onda del ultravioleta sobre la piel de las lagartijas produce la síntesis de vitamina D (al igual que ocurre en los mamíferos, hombre incluido). Esta vitamina es esencial para un correcto metabolismo del calcio. Al aplicárseles la radiación requerida en el laboratorio de herpetología del Museo Americano, Carol R. Townsend y el autor descubrieron, además, que las diversas especies de *Cnemidophorus* precisaban distintas dosis de radiación.

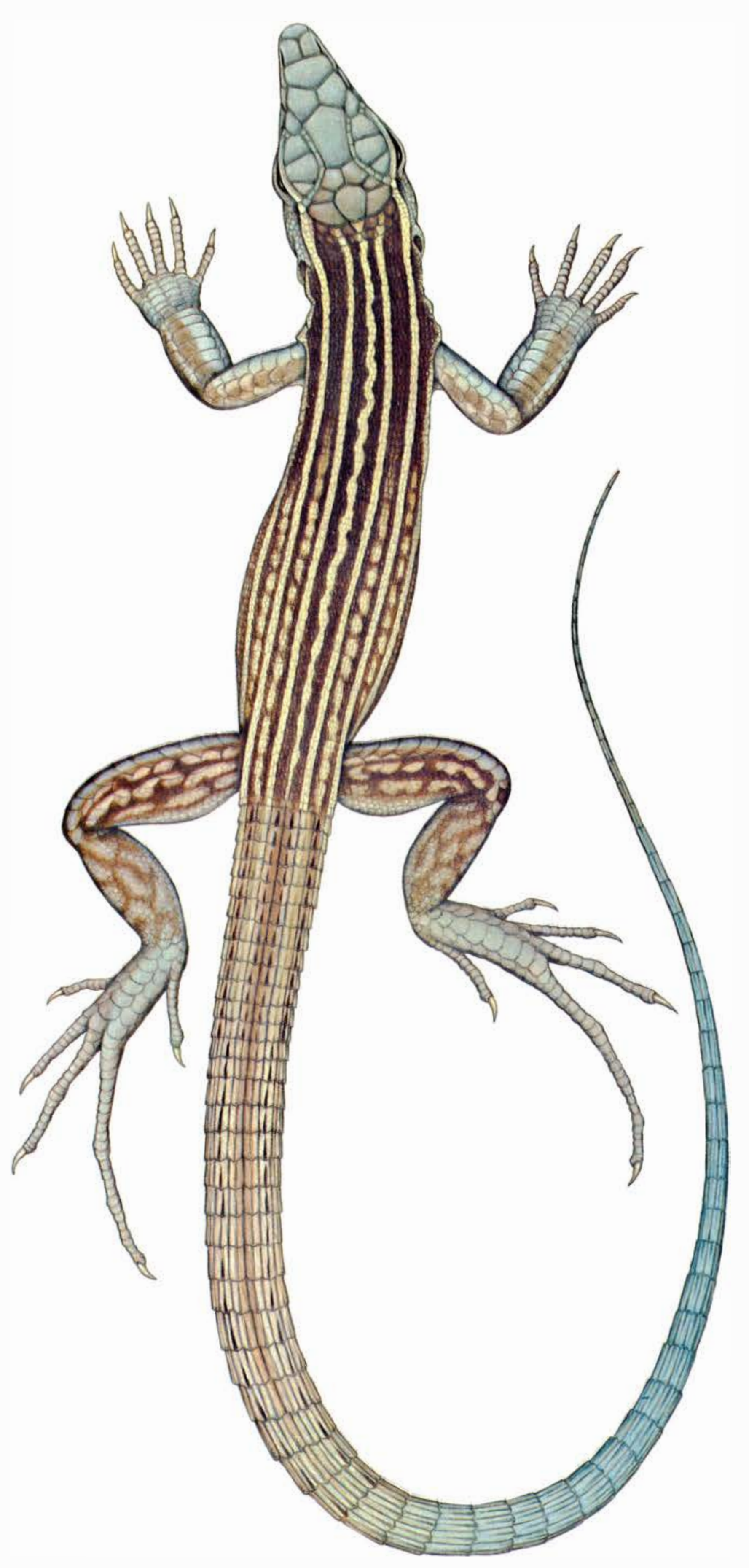
Una vez resueltos esos detalles, la cría de muchas especies de *Cnemidophorus* unisexuales en cautividad se convirtió en mera rutina. En nuestro laboratorio hay estirpes unisexuales que no sólo han alcanzado la meta inicial de una segunda generación, sino genera-

ciones sucesivas hasta la séptima, con descendientes que se cuentan por centenares y sin un solo macho entre ellos.

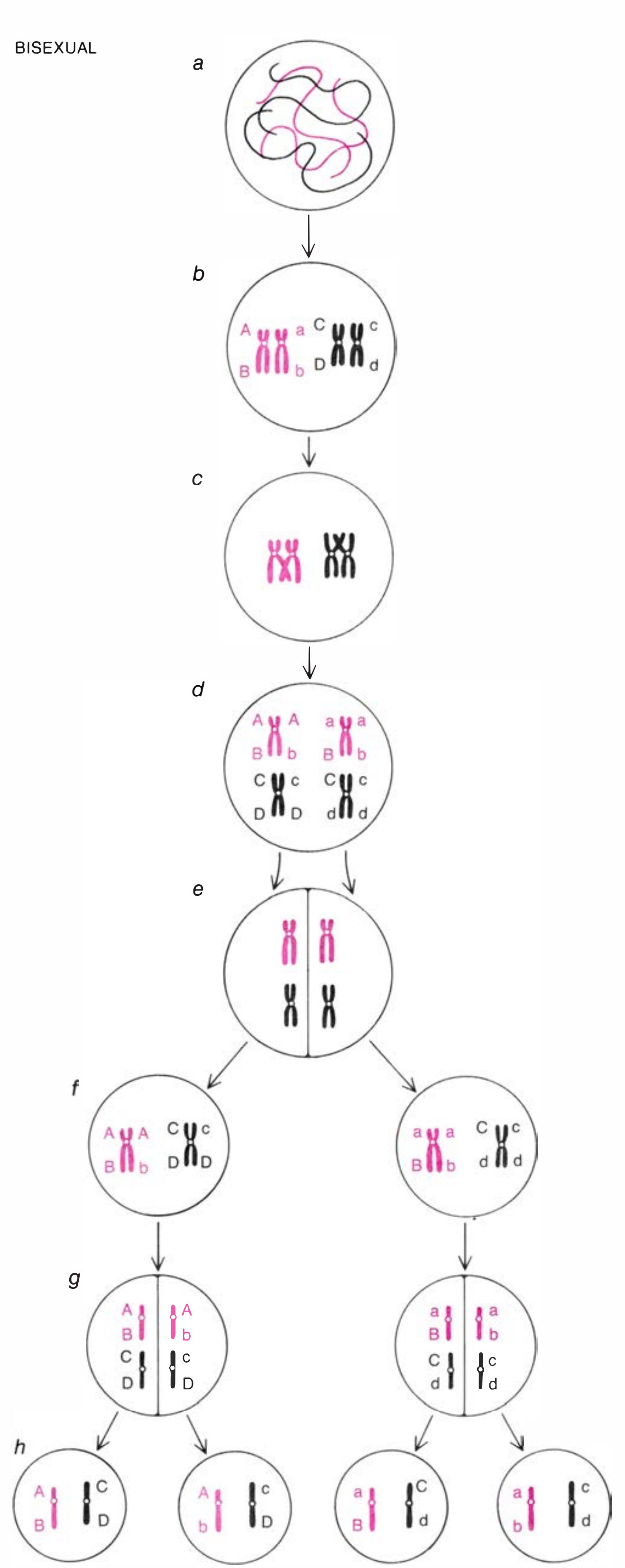
A finales de la década de 1970, las pruebas indicaban de forma incontrovertible que las especies de *Cnemidophorus* que sólo tenían hembras se reproducían mediante partenogénesis estricta, es decir, en ausencia total de espermatozoides. No era una conclusión para tomársela a la ligera, pues se trataba del primer caso conocido de especies partenogenéticas en los vertebrados. Las pruebas más claras de reproducción partenogenética se obtuvieron al comparar las pautas de desarrollo en hembras unisexuales y en lagartijas bisexuales de ambos sexos, con referencia particular a su aparato reproductor. Laurence M. Hardy, de la Universidad estatal de Louisiana en Shreveport, realizó la mayor parte de este trabajo, con descendientes de nuestra segunda generación.

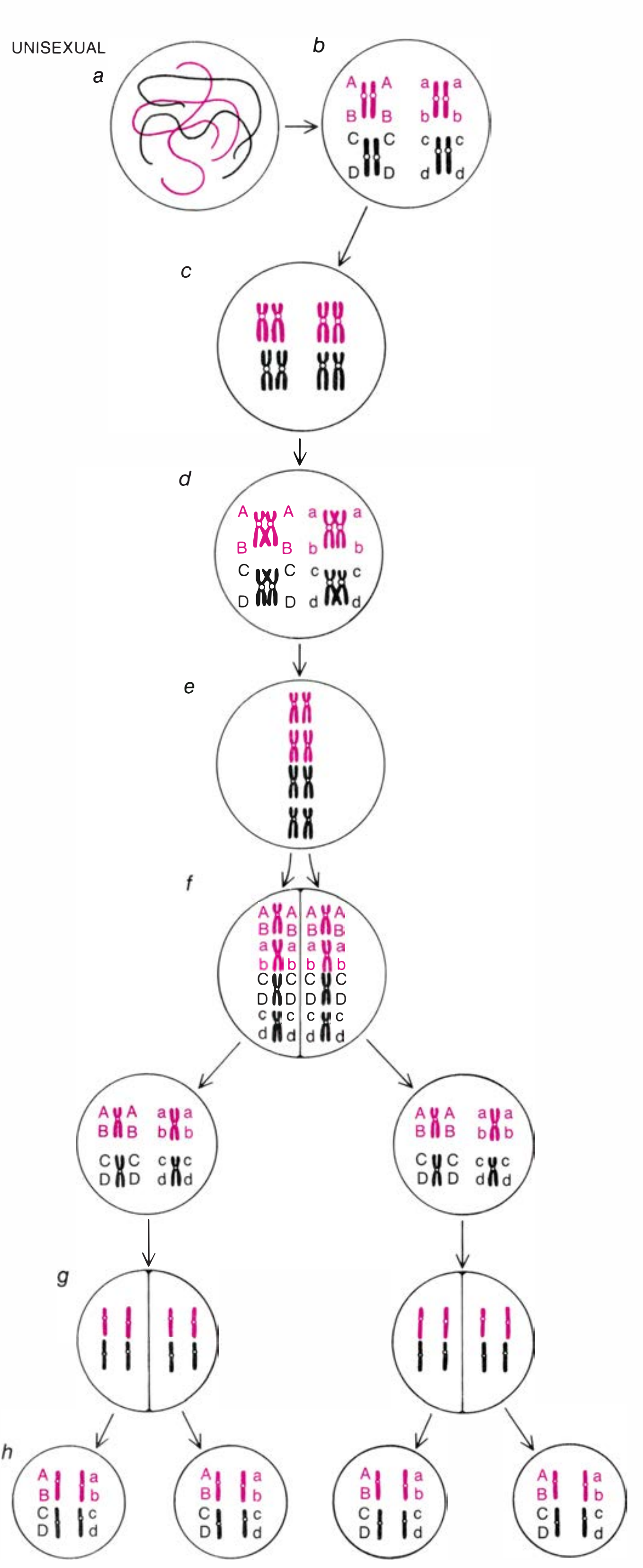
Hardy encontró que el aparato reproductor de las lagartijas unisexuales era idéntico en los rasgos principales al de las hembras de las especies bisexuales. No halló señal alguna de esperma en ninguno de los *Cnemidophorus* unisexuales, lo que prestaba un apoyo concluyente a nuestra hipótesis de que se reproducían por partenogénesis estricta.

En los años que precedieron al establecimiento de colonias de *Cnemidophorus* en los laboratorios, varios biólogos, entre ellos Zweifel, T. Paul Maslin, de la Universidad de Colorado en Boulder, y William Neaves, de la Facultad de Medicina de Harvard, estudiaron lagartijas de cola de látigo unisexuales, capturadas en el campo, para contrastar la hipótesis de que sus poblaciones eran clones, es decir, se componían de individuos genéticamente idénticos emparentados por una ascendencia común. Vale la pena resumir sus conclusiones. En primer lugar, en lo que respecta a características tales como el número de escamas epidérmicas localizadas hacia la mitad del cuerpo de la lagartija, las especies unisexuales presentaban menos variabilidad que las especies bisexuales vecinas. En segundo lugar, las células somáticas (es decir, del cuerpo) de las lagartijas de las especies unisexuales contenían cromosomas marcadores que indicaban la presencia de dos genomas (dotaciones completas de genes) distintos; vale decir, los cromosomas eran heterocigotos. Además, cuando se producía esta circunstancia, el estado heterocigoto parecía fijado en la población, pues todos los individuos eran iguales. Por su parte, los cromosomas de las lagar-



1. LAGARTIJA DE COLA DE LATIGO del Nuevo Mundo, del género *Cnemidophorus*, dibujada a tamaño doble del natural. La mayoría de especies de *Cnemidophorus* son bisexuales; su reproducción precisa de la fecundación de los óvulos de la hembra por parte de los espermatozoides del macho. Esta hembra pertenece a una de las especies unisexuales: *C. neomexicanus*. Sus óvulos se desarrollan en ausencia total de espermatozoides y sus descendientes son también hembras de reproducción partenogenética.





2. PROCESO DE LA MEIOSIS en los *Cnemidophori* bisexuales, comparado con el mismo proceso en las especies unisexuales. En la parte izquierda (*a*) se ilustran esquemáticamente dos parejas de cromosomas (*negro y color*) en una célula ovárica diploide, es decir, que posee dos dotaciones cromosómicas. Los cromosomas se duplican y se aparean (*b*); luego pueden entrecruzarse (*c, d*), proceso que (como ocurre aquí en *d*) determina que la combinación de genes de ciertos cromosomas (*indicados mediante las letras* A, a, B, b, C, c, D y d) difiera de la inicial. Sigue luego la primera división celular (*e*), en la que un cromosoma de cada dotación se sitúa en cada nueva célula (*f*). Una segunda división celular (*g*) origina cuatro nuevas células; una de ellas se transformará en óvulo funcional. Estas células son haploides, es decir, tienen una sola dotación de cromosomas (*h*). La unión con un espermatozoide haploide, o sea, la fecundación, hará que la célula ovular retorne al estado diploide viable. En el lado derecho de la ilustración aparece el proceso que tiene lugar en la célula de una lagartija unisexual. El inicio (*a*) es similar: una célula diploide con dos pares de cromosomas. Pero aquí los cromosomas se duplican y se separan (*b*). Luego se duplican de nuevo y se aparean (*c*), ahora en una célula tetraploide. Los nuevos cromosomas se aparean con aquellos de los que proceden. El entrecruzamiento (*d*) va seguido por la primera división de la célula ovárica tetraploide (*e*), en la que un cromosoma de cada dotación duplicada se dirige a cada nueva célula (*f*). La división siguiente (*g*) origina cuatro nuevas células, una de las cuales se transforma en un óvulo funcional. Estas células son diploides (*h*) y genéticamente idénticas a la célula ovárica original. El desarrollo de estos óvulos procederá sin intervención alguna de espermatozoides.

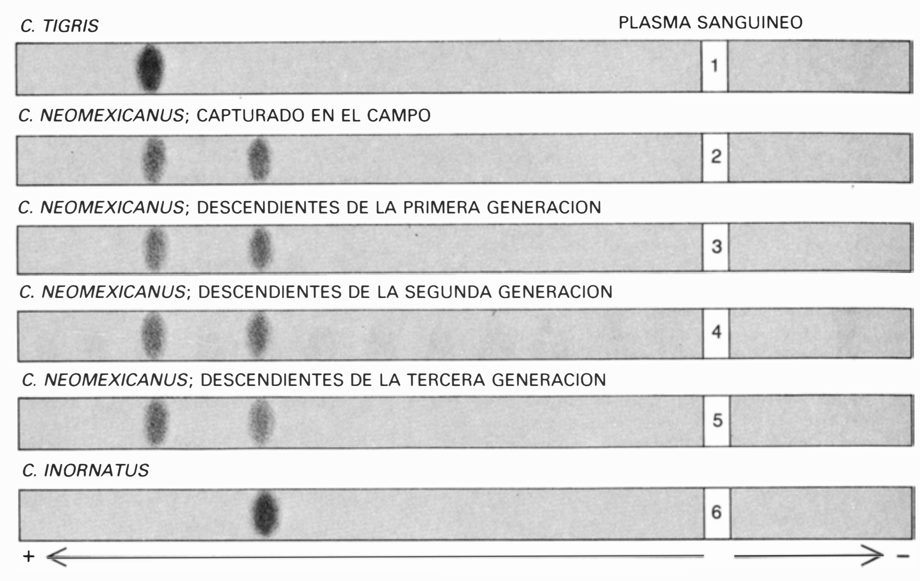
tijas de las especies bisexuales eran homocigotos: los animales tenían dos genomas similares.

En tercer lugar, las células somáticas de las lagartijas de muchas especies unisexuales presentaban un número cromosómico insólitamente elevado. En vez del número diploide normal, característico de las especies bisexuales, el suyo era triploide. En cuarto lugar, al separar por electroforesis las proteínas de las especies unisexuales se detectaba un elevado nivel de heterocigosis, aparentemente fijada, mientras que en las proteínas de las especies bisexuales no se daba ese caso. Finalmente, cuando se trasplantaba un fragmento de piel de una lagartija a otra de la misma especie unisexual, el injerto solía prender, prueba de que entre los individuos de la misma especie existía una buena compatibilidad hística. Con las especies de lagartijas de cola de látigo bisexuales ocurría exactamente lo contrario. Todos estos resultados encajaban en la hipótesis de que los reptiles unisexuales eran clones.

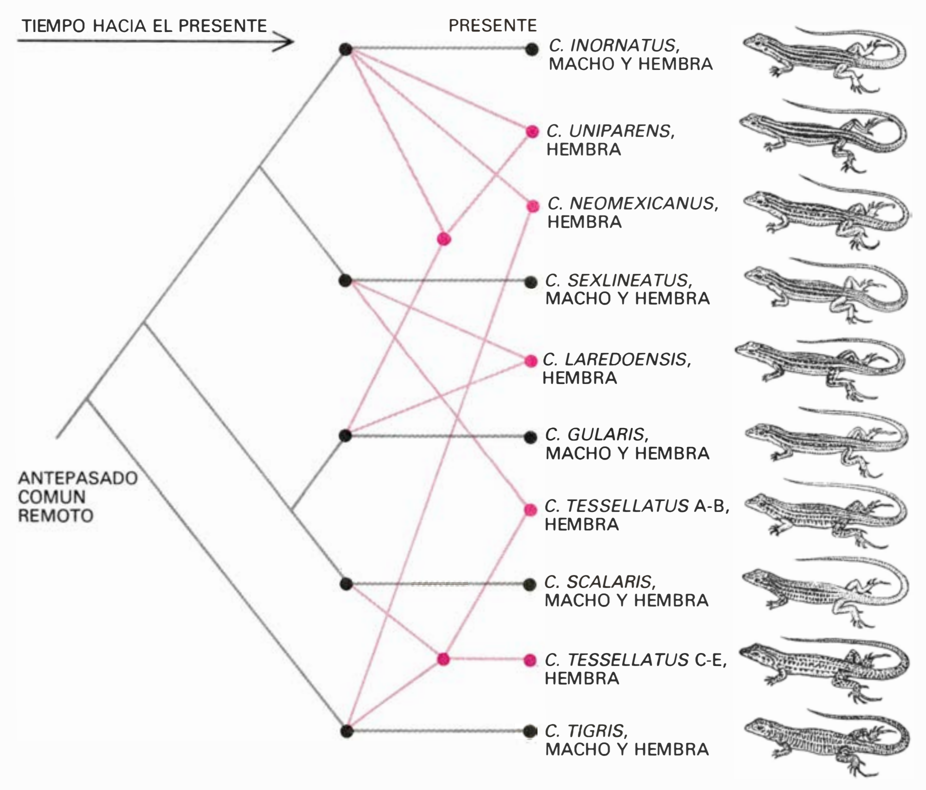
Las colonias de laboratorio de los *Cnemidophorus* unisexuales han permitido reunir datos sobre los clones que no se hubieran podido recabar de lagartijas capturadas en la naturaleza. Por ejemplo, en nuestro laboratorio, los huevos y las lagartijas recién eclosionadas se mantienen de tal modo que pueda siempre establecerse la identidad y la ascendencia de cualquier lagartija. Lo que facilita rastrear la genealogía de un individuo y de sus parientes hasta la hembra original que se cazó en el campo. A través de una estirpe unisexual podemos investigar la morfología externa, estados cromosómicos y pautas electroforéticas de las proteínas, entre otros.

De los estudios efectuados con nuestras colonias de laboratorio se deriva lo siguiente: en primer lugar, morfología, color, pautas de éste y forma y tamaño de las escamas de las lagartijas unisexuales son casi siempre idénticos a los de la madre, con algunas variaciones inesperadas que hoy son objeto de investigación para averiguar si tienen fundamento hereditario o se deben al medio en que crecen. En segundo lugar, tanto si la especie de la madre posee un juego diploide de cromosomas como si lo posee triploide, ésta produce descendientes con cromosomas idénticos a los suyos propios.

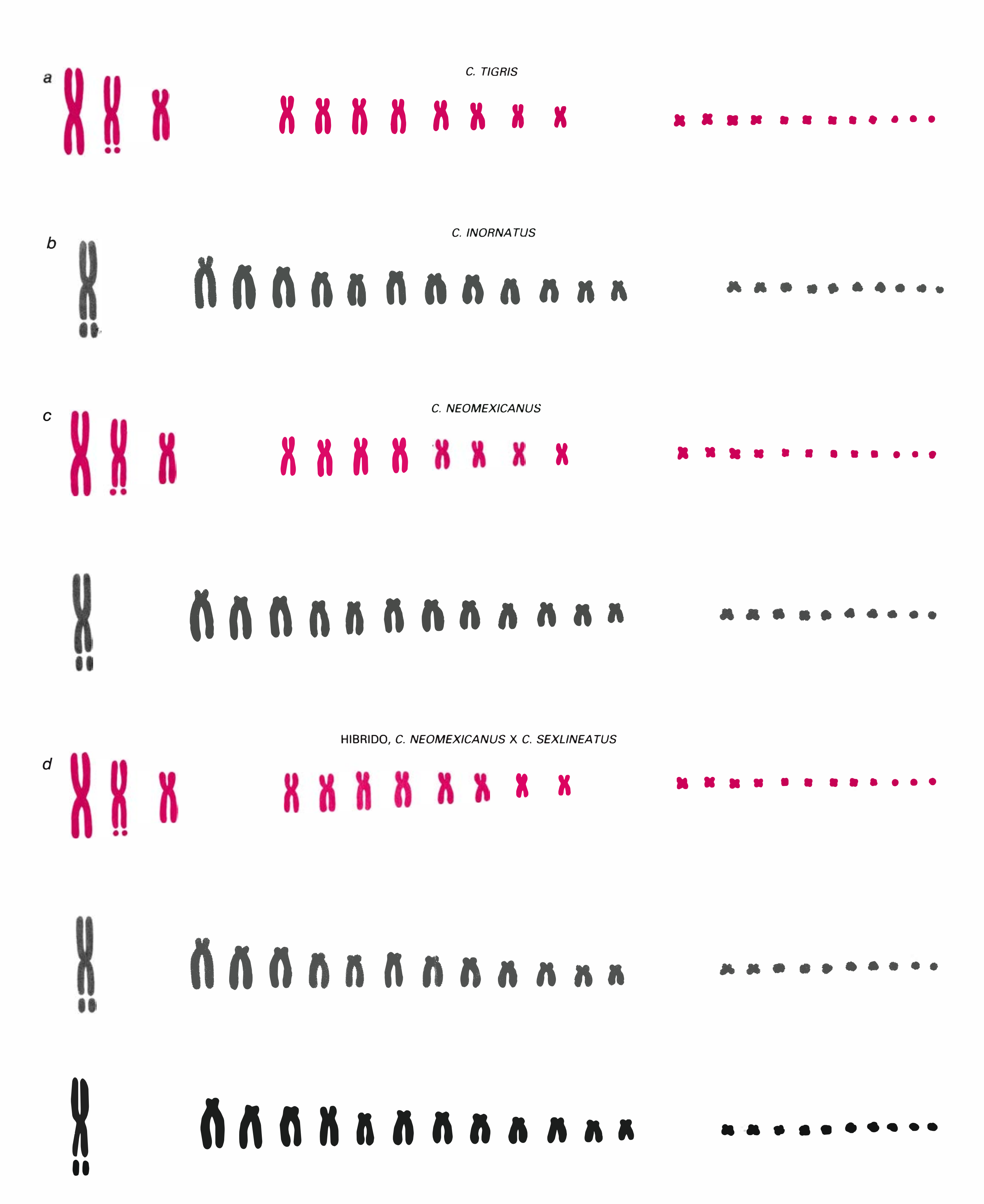
Un tercer hallazgo de laboratorio, en el campo de la genética bioquímica, procede de los estudios de Herbert



3. PROTEINAS SANGUINEAS de tres especies de la lagartija *Cnemidophorus*; se analizaron por electroforesis en gel, cuyo resultado, esquematizado, se ofrece aquí. Se introdujeron muestras de plasma sanguíneo mezcladas con un isótopo radiactivo del hierro en las seis ranuras. Después de la electroforesis, el hierro radiactivo impresionó una hoja de película fotográfica dispuesta sobre el gel, señalando las posiciones de la transferrina, una proteína que fija hierro. Todas estas moléculas proteínicas se habían desplazado hacia el polo positivo del gel. Las dos manchas más oscuras que quedan frente a las ranunas *1* y *6* son, respectivamente, las moléculas rápidas de la transferrina de la sangre de *C. tigris* y las moléculas, más lentas, de una transferrina estructuralmente distinta de la sangre de *C. inornatus*. Las ocho manchas más claras marcan el movimiento de moléculas de las mismas proteínas de la sangre de la especie unisexual *C. neomexicanus*. La muestra de la ranura 2 procede de un ejemplar de esta especie capturado en la naturaleza. Las muestras de las restantes ranuras proceden de sus descendientes de generaciones sucesivas. Su desplazamiento idéntico demuestra herencia clónica: las proteínas, y por tanto los genes que las codifican, son los mismos de una generación a la siguiente. Además, la especie unisexual *C. neomexicanus* expresa los genes de las dos transferrinas distintas presentes en sus especies ancestrales. Muchos otros polipéptidos revelan que sus genes se hallan en estado de heterocigosis en las especies partenogenéticas.

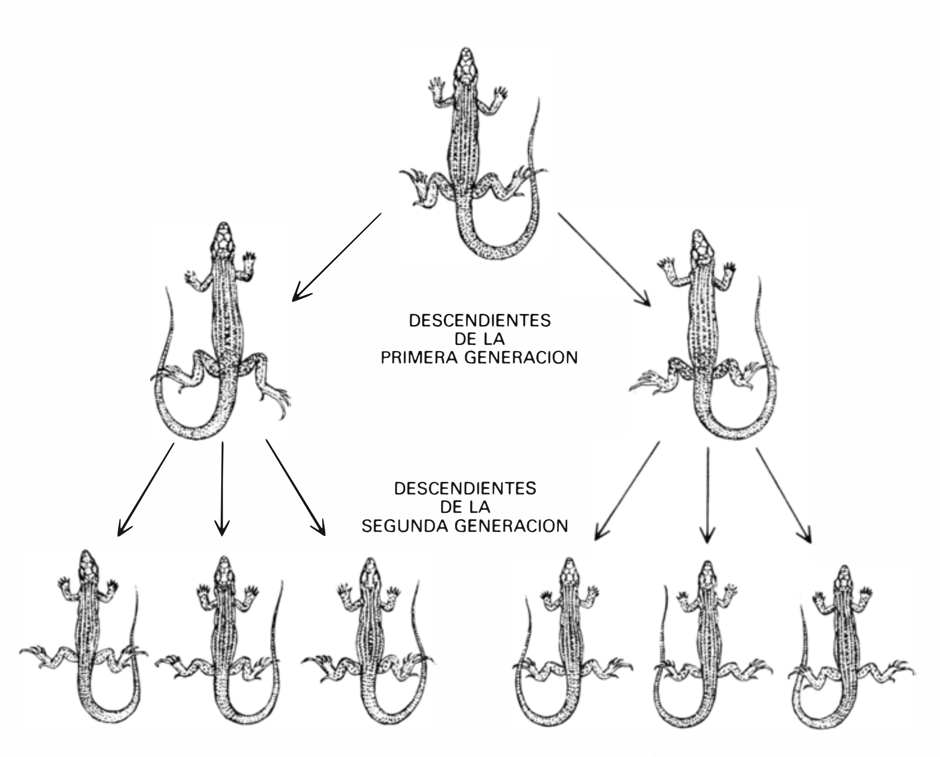


4. RELACIONES FILOGENETICAS entre nueve especies de *Cnemidophorus*, muchas de ellas unisexuales. Las especies bisexuales (*negro*) se emparentan de acuerdo con la teoría de la especiación: divergencia a partir de antepasados comunes. Aquí la mayor divergencia se da entre *C. inornatus* (*arriba*) y *C. tigris* (*abajo*). Las especies unisexuales (*color*), derivadas de las anteriores mediante hibridación, presentan una pauta filogenética que guarda más parecido con una red. Las especies unisexuales diploides, como *C. neomexicanus* y *C. laredoensis*, han surgido de una hibridación única, pero las triploides, como *C. uniparens* y *C. tesselatus* A-B, tienen ascendencias que incluyen hibridaciones y a veces tres especies bisexuales.



5. CROMOSOMAS de varias especies de lagartijas de cola de látigo, ilustrados esquemáticamente. Arriba (*a*) los 23 cromosomas de una célula germinal de *Cnemidophorus tigris*, su número haploide. Bajo éstos (*b*) se hallan los 23 cromosomas de una célula germinal haploide de *C. inornatus*. Siguen los 46 de una célula somática típica de la especie unisexual *C. neomexicanus*, la dotación diploide usual (*c*). Su semejanza con las dos dotaciones de cromosomas superiores concuerda con la hipótesis de que el origen de *C. neomexicanus* fue un cruzamiento entre las especies bisexuales *C. tigris* y *C. inornatus*. Como se ilustra en la figura 2, esta combinación de cromosomas persiste, por clonación, en las sucesivas generaciones partenogenéticas de *C. neomexicanus*. Si una de estas hembras unisexuales se cruzara con el macho de una especie bisexual, se produciría un híbrido con tres dotaciones de cromosomas (*d*). El juego triploide de 69 cromosomas procede de una célula somática obtenida de un cruzamiento de laboratorio entre una hembra de *C. neomexicanus* y un macho de *C. sexlineatus*, bisexual. Si un triploide se cruzara con un macho normal, las células somáticas de los descendientes serían tetraploides.

Dessauer, del Hospital Clínico de la Universidad estatal de Louisiana en Nueva Orleans. Se ha ocupado de las proteínas producidas por más de 30 genes y ha descubierto que, generación tras generación, las proteínas de los descendientes de *Cnemidophorus* unisexuales presentan las mismas pautas electroforéticas que las de la madre. Se expresan los mismos genes y, todas las combinaciones genéticas, incluyendo los estados heterocigotos, están fijadas. Los experimentos de laboratorio refuerzan, por tanto, las conclusiones de los estudios de campo.



6. LAGARTIJAS DE COLA DE LATIGO unisexuales de la especie *Cnemidophorus exsanguis*, de las que se han criado en el laboratorio siete generaciones partenogenéticas sucesivas. Aquí se ilustran el progenitor inicial y algunos descendientes de la primera y de la segunda generaciones. Todos los descendientes son hembras genéticamente idénticas entre sí y a su progenitor. Los seis descendientes de la segunda generación no presentan la coloración de los demás: las marcas de los adultos aparecen más tarde.

Observaciones sobre el proceso que, al clonarse, siguen los óvulos de las lagartijas partenogenéticas indican que los cromosomas se duplican dentro de las células ováricas destinadas a convertirse en óvulos. Así, estas células reproductoras especializadas se hacen poliploides. En las especies diploides se convierten en tetraploides; en las triploides, hexaploides. Los recuentos cromosómicos, sumados a las pruebas de herencia clónica, sugieren que, al aparearse para producir óvulos, los cromosomas duplicados se emparejan con los cromosomas de los que son copia exacta. El intercambio de ADN en el entrecruzamiento entre cromosomas apareados no comporta consecuencias genéticas porque afecta a hebras idénticas de ADN. El entrecruzamiento va seguido de las dos divisiones celulares típicas de la maduración de los óvulos, de modo que el producto final es un óvulo que contiene los mismos cromosomas y la misma combinación de genes que se daban en la célula ovárica original antes de la duplicación de los cromosomas. El óvulo maduro, por tanto, está genéticamente completo, en lugar de ser haploide, como ocurre en las hembras de las especies bisexuales, y da comienzo al desarrollo obviando la fecundación.

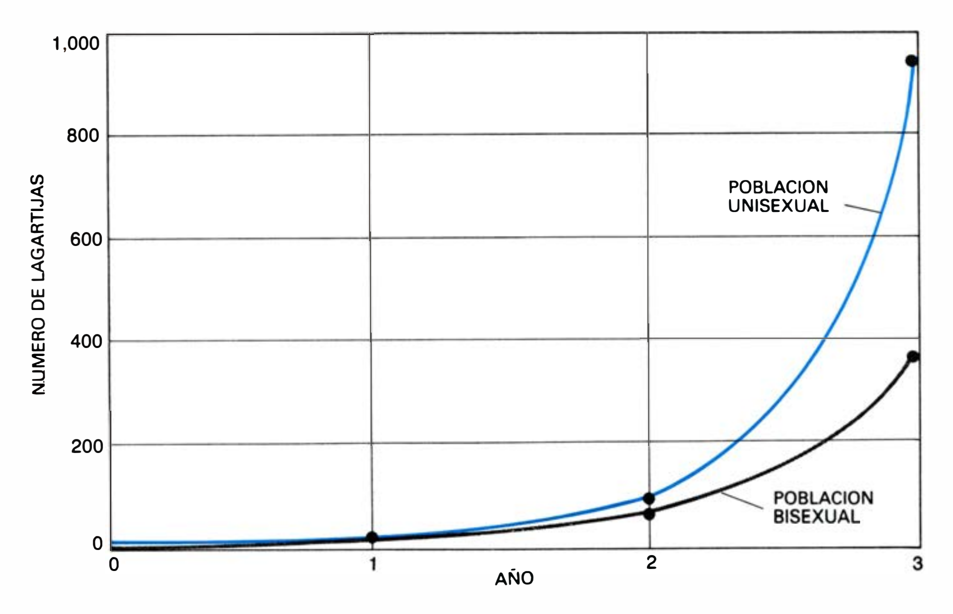
¿Cómo surgió en la naturaleza esta peculiar partenogénesis? El género de lagartijas de cola de látigo *Cnemidophorus*, confinado al Nuevo Mundo, comprende unas 40 especies. La mayoría de ellas son bisexuales, pero quizás una docena sean unisexuales. La mayoría de especies unisexuales se encuentran en una región de los Estados Unidos sudoccidentales y de México septentrional donde las comunidades vegetales han sufrido cambios y mezclas en los períodos de variaciones climáticas de la época geológica reciente. En algunos lugares, una o dos especies bisexuales de lagartijas de cola de látigo comparten su área de distribución con una especie unisexual.

En estas circunstancias, se produce a veces la inseminación de una hembra partenogenética por un macho de una especie bisexual, y de ello resulta una descendencia híbrida. Los descendientes tienen números cromosómicos poliploides: hembras unisexuales diploides y triploides producen, respectivamente, descendientes híbridos triploides y tetraploides. Dessauer ha encontrado que todos los genes que la electroforesis de proteínas señala como característicos de cada progenitor son funcionales. Los híbridos exhiben más rasgos de la madre (porque ésta ha aportado dos o tres juegos de cromosomas) que del padre (que sólo ha aportado un juego). Si el óvulo resulta fecundado por un espermatozoide que porte un cromosoma *X*, el híbrido será hembra; si lo fecunda un espermatozoide portador de un cromosoma *Y*, el híbrido será macho.

Las características de estos híbridos sugieren que, en la naturaleza, las especies de *Cnemidophorus* unisexuales y triploides se clonaron a partir de descendientes femeninos híbridos de la inseminación de una hembra partenogenética diploide por un macho de una especie distinta. Pero, ¿cómo surgió la hembra partenogenética diploide? Charles H. Lowe, Jr., de la Universidad de Arizona, y John W. Wright, que en la actualidad se halla en el Museo de Historia Natural del condado de Los Angeles, dieron, en 1966, con la respuesta. Una de las especies que examinaron con más detalle fue la lagartija unisexual de Nuevo México *C. neomexicanus*. La especie se encuentra a lo largo de las corrientes de agua del sistema hidrográfico de Río Grande y en hábitats colindantes de desierto y pradera, es decir, zonas de transición entre el desierto y la pradera. Una especie bisexual, *Cnemidophorus inornatus*, ocupa un hábitat de pradera que comparte marginalmente con la unisexual *C. neomexicanus*. Otra especie bisexual, *C. tigris*, ocupa un hábitat de desierto que también comparte marginalmente con *C. neomexicanus*.

Lowe y Wright observaron que, en rasgos tan pertinentes como color, pauta del mismo, forma de las escamas, cromosomas y hábitats preferidos, la idiosincrasia de *C. neomexicanus* semeja la de un híbrido de la primera generación resultante del apareamiento de *C. inornatus* y *C. tigris*. Mediante clonación partenogenética, las poblaciones de *C. neomexicanus* han conservado su identidad, en tanto que híbridos de la primera generación, en las sucesivas generaciones.

Lowe y Wright sugirieron que, en algún momento del pasado, las especies



7. SU ELEVADO POTENCIAL REPRODUCTOR constituye un carácter ventajoso de los lagartos unisexuales. Este gráfico compara el crecimiento de una población de lagartijas de cola de látigo unisexuales (*color*) con el de una población de individuos bisexuales a lo largo de un período de tres años. Los datos se basan en la producción de huevos por un ejemplar de *Cnemidophorus exsanguis* criado en el laboratorio. El punto en el año 0 del gráfico representa la eclosión de un huevo. Para la especie bisexual se precisa la presencia breve de un macho para la fecundación en el año 1. Al final del primer año, y suponiendo un 100 por cien de eclosión y supervivencia, ambas poblaciones estarían formadas por nueve individuos. Sin embargo, al final del tercer año las lagartijas unisexuales doblarían en número a las bisexuales, pues todas las unisexuales ponen huevos, mientras que sólo se da ese caso en la mitad de las bisexuales.

de pradera de *Cnemidophorus* se cruzaron con las especies de desierto, produciendo híbridos como *C. neomexicanus*. Igual que ocurre con la mayoría de híbridos interespecíficos, los cruzamientos iniciales produjeron probablemente machos y hembras, la mayoría de ellos estériles. Al mismo tiempo, es probable que los híbridos compitieran con éxito con las formas no híbridas por los hábitats mixtos de desierto y pradera. Los machos híbridos de la primera generación desaparecerían, pero hembras capaces de clonación partenogenética habrían perpetuado su grupo.

Suele presentarse la esterilidad de los híbridos interespecíficos como ejemplo de actuación de la selección natural. Ello equivale a decir que los individuos de una especie que se cruzan sólo con otros de su propia especie producen los descendientes más fértiles, mientras que la hibridación, con su esterilidad inherente, parecería no tener ninguna ventaja adaptativa. Sin embargo, en el caso de las lagartijas unisexuales del género *Cnemidophorus*, los animales han continuado perpetuándose mediante clonación, a pesar de su origen híbrido. ¿Gozan estas lagartijas partenogenéticas de alguna ventaja adaptativa?

Desde luego, una ventaja inmediata es el hecho de que la partenogénesis anula la desventaja que usualmente impone la esterilidad híbrida. La herencia clónica parece asimismo mantener la ventaja que a veces se denomina vigor híbrido, esto es, mediante la clonación, los animales no desbaratan una nueva combinación de genes que ha resultado adecuada en un ambiente determinado.

La tercera ventaja arranca del potencial reproductor de las lagartijas: puesto que toda la población está compuesta por hembras que dan origen a más hembras, las especies unisexuales pueden reproducirse en mayor número que las bisexuales, de las que la mitad de los individuos (los machos) no pueden poner huevos. También gozan de alguna ventaja al colonizar nuevas localidades y quizás incluso nuevos hábitats. Por ejemplo, en el caso de dispersión al azar, la supervivencia de una sola lagartija partenogenética basta para establecer una nueva población en una zona antes deshabitada. En algunos casos estas últimas ventajas podrían conducir al establecimiento de partenogénesis mediante mutación, y no mediante hibridación; este pudiera ser el caso de la especie sudamericana *Cnemidophorus lemniscatus*, estudiada por P. E. Vanzolini, de la Universidad brasileña de São Paulo.

Sin embargo, y al menos en principio, no todo son ventajas para una especie partenogenética. La herencia clónica es ventajosa en condiciones ambientales fijas porque conserva las combinaciones génicas que favorecen la adaptación a dicho medio. Pero las condiciones ambientales raramente son estables. Cuando cambian deprisa, como ocurre con frecuencia, las especies bisexuales tienen más posibilidades que las unisexuales de producir algún descendiente diferente que se adapte a la nueva situación. Por ello, la probabilidad de supervivencia a largo plazo puede ser menos favorable para las especies que se reproducen mediante clonación que para las que se reproducen sexualmente según las leyes de la herencia mendeliana.

Aún así, no todas las cartas evolutivas juegan en contra de las especies partenogenéticas. Más que desaparecer, los hábitats suelen cambiar de lugar. En tales circunstancias, las lagartijas clónicas sólo necesitan desplazarse con su hábitat para permanecer dentro del mismo. Además, si en una estirpe unisexual ocurriera cualquier mutación ventajosa, la clonación aseguraría su transmisión a las generaciones sucesivas. La clonación puede conferir incluso cierta protección frente a las mutaciones desventajosas, en particular si son recesivas y se producen en individuos triploides.

Hace 25 años, algunos biólogos plantearon valientemente la posibilidad de que existieran especies de saurios unisexuales. Por aquella época nadie hubiera vaticinado que tales investigaciones conducirían a los sorprendentes descubrimientos de la clonación partenogenética, la poliploidía y el origen de nuevas especies de vertebrados por medio de la hibridación. Hoy, los investigadores interesados en el tema, además de plantearse nuevas preguntas y diseñar nuevos experimentos, disponen de estos especializados organismos para ensayos no imaginables hace escasos años. Entre las posibilidades que a uno se le ocurren se cuentan la de obtener un mejor conocimiento del papel de los espermatozoides en la fecundación; esclarecer por qué a algunos animales les favorece la posesión de múltiples copias de sus genes y a otros no; estudiar los mecanismos de desviación en el desarrollo embrionario; producir animales clónicos de composición genética conocida para la experimentación biológica e incluso la de inducir la clonación en especies normalmente bisexuales para aumentar la productividad ganadera. Si los reptiles unisexuales contribuyen al progreso en alguno de estos campos, o en todos, habrá sido a partir de unas pocas y sorprendentes observaciones sobre una forma de vida salvaje que nadie apreciaba en su valor.

Control de la síntesis de ribosomas

La partícula que traduce el ARN en proteínas contiene tres moléculas de ARN y 52 proteínas: ¿Cómo se adapta el ensamblaje de los ribosomas, a partir de estos componentes, a las necesidades cambiantes de la célula?

Masayasu Nomura

Las células sintetizan sus proteínas en los ribosomas, donde los aminoácidos se añaden secuencialmente a la cadena de proteína que se esté formando. Una vez terminada, la cadena se pliega, resultando una molécula que puede desempeñar funciones estructurales o enzimáticas; en este segundo caso, como catalizadora de una reacción metabólica específica. Todas las funciones metabólicas requieren el concurso de enzimas; por tanto, para que una célula crezca y se divida más rápidamente debe incrementar su capacidad de síntesis de proteínas. El ribosoma tiene, sin embargo, una capacidad de trabajo limitada. En la bacteria *Escherichia coli*, por ejemplo, la velocidad de síntesis de proteínas es de 15 aminoácidos por segundo, a una temperatura de 37 grados Celsius. Si la temperatura permanece constante, y el medio de cultivo se enriquece con nutrientes, la tasa de división de las bacterias aumenta. Por ser limitada la capacidad de los ribosomas, si la célula quiere incrementar su tasa de síntesis de proteínas deberá fabricar más ribosomas. La fabricación de ribosomas es cara, desde el punto de vista energético, y tener un exceso de ellos constituye un despilfarro. La bacteria, por tanto, debe poder ajustar su producción de ribosomas con gran precisión.

No resulta fácil controlar la tasa de síntesis de ribosomas, habida cuenta de la complejidad del orgánulo. Cada ribosoma contiene 52 proteínas diferentes y tres moléculas de ARN distintas. La información para fabricar las proteínas y ARN ribosómicos (ARNr) se halla en el ADN bacteriano; el proceso de síntesis de proteínas comienza cuando dicha información se transcribe en una molécula de ARN mensajero (ARNm). Posteriormente, el ARNm se traduce en proteína, operación que se realiza en los ribosomas. Los ARN ribosómicos se transcriben directamente del cromosoma. Para conseguir un ritmo adecuado de síntesis de ribosomas, la producción de todos sus componentes debe realizarse en perfecta sincronía, de la misma forma que, para que un coche se fabrique con la máxima eficacia, sus componentes deben prepararse en las cantidades justas y distribuirse por la cadena de montaje en el momento apropiado. El control puede ejercerse en dos niveles, al menos: en la transcripción de los genes responsables de los 55 componentes ribosómicos o en la traducción de los ARNm correspondientes a las 52 proteínas.

Como método de control más lógico cabría pensar, en un principio, en la regulación directa de la transcripción. A comienzos de la década de 1970, cuando se iniciaron los trabajos sobre control de la síntesis ribosómica, éste era el modelo de control genético en boga. No obstante, pronto se comprobó que un control directo no explicaba satisfactoriamente todos los aspectos conocidos del proceso de ensamblaje de ribosomas. Trabajos posteriores, que se revisarán en este artículo, mostraban que el control de la síntesis de ribosomas se basaba fundamentalmente en la existencia de dos mecanismos de retrorregulación negativa, o retroinhibición. Mientras haya moléculas de ARN ribosómico disponibles, las proteínas ribosómicas se unirán a ellas y se irán ensamblando ribosomas. Cuando se agote el ARNr, proteínas ribosómicas actuarán de "represoras" de la traducción, uniéndose a los ARN mensajeros e interrumpiendo la síntesis de todas las proteínas ribosómicas. Ya tenemos el primer mecanismo de retroinhibición.

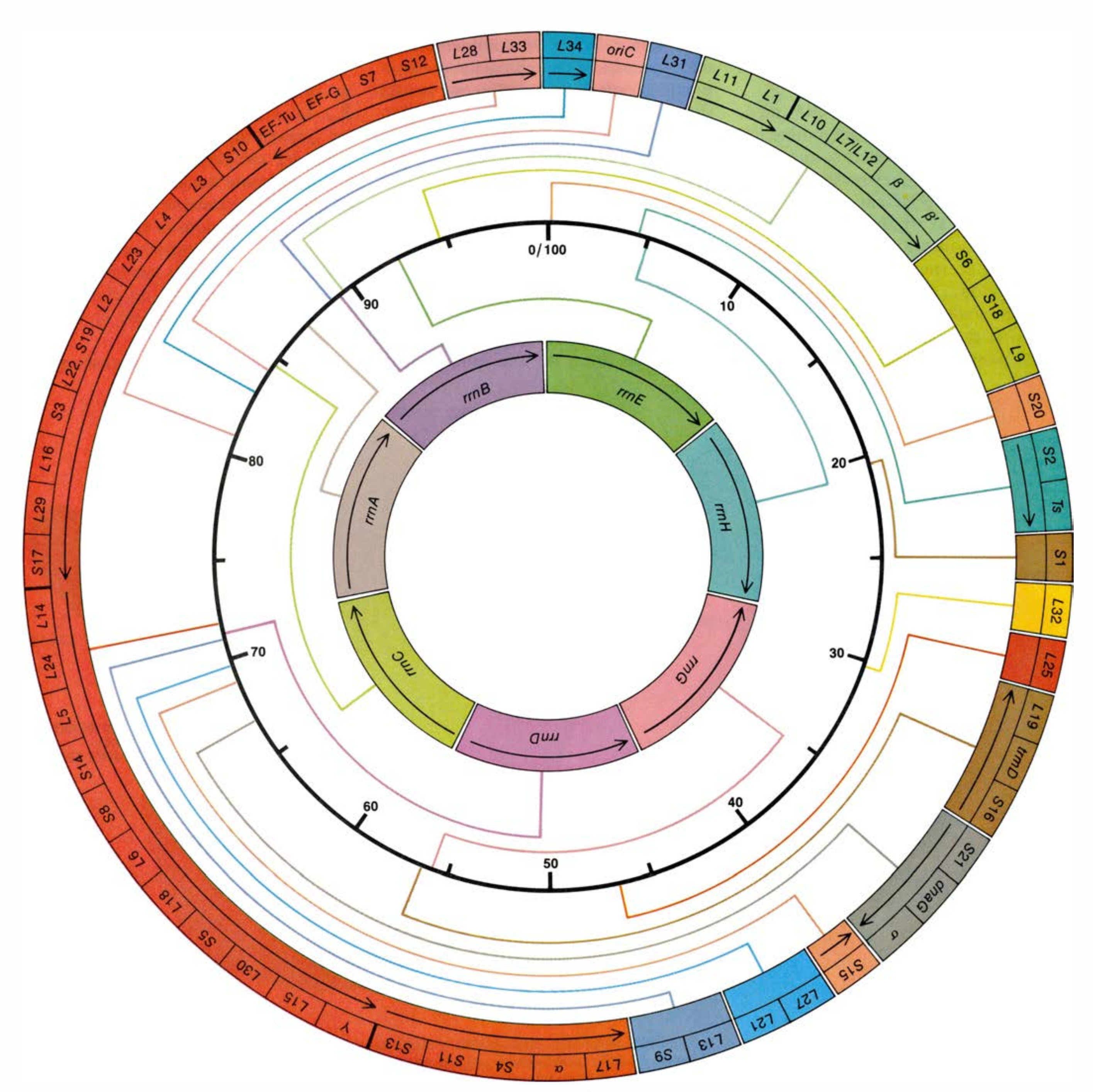
El segundo modula la transcripción, entre otros, de los genes de los ARN ribosómicos. El control lo ejercen aquí los ribosomas "libres", es decir, los que no están ocupados en sintetizar proteínas. Cuando se acumulan ribosomas libres, éstos "desconectan" los genes de los ARNr. Si el medio se enriquece en nutrientes, los ribosomas libres que haya deben ahora ocuparse de la síntesis de proteínas y los genes de ARNr se activan de nuevo. La acción conjunta e integrada de los dos mecanismos reguladores permite que la bacteria ajuste con gran precisión el ritmo de síntesis de ribosomas. Al igual que ha ocurrido con otras muchas investigaciones científicas fundamentales, el descubrimiento de este sistema de control genético ha desplegado un abanico de posibilidades. La represión de la traducción mediante un mecanismo de retroinhibición podría explicar el control de la síntesis de otras estructuras bacterianas complejas. No sólo eso; los resultados obtenidos con *E. coli* quizá podrían extrapolarse al control de la síntesis de ribosomas en las células, más complejas, de los organismos multicelulares, hombre incluido. Por otra parte, el desentrañamiento de un sistema de regulación biológica tan intrincado y elegante produce una satisfacción estética considerable, lo que constituye una porción nada despreciable de los frutos de un trabajo como ése.

La estructura, función y ensamblaje del ribosoma están tan estrechamente ligados a la expresión de la información genética que se hace necesario conocer las bases del código genético para comprender cómo se regula la síntesis de ribosomas. El ADN, material genético primario, es una cadena formada por cuatro tipos de bases nucleotídicas: adenina, citosina, guanina y timina (*A*, *C*, *G* y *T*, respectivamente). Por sus características estructurales, *A* normalmente forma enlaces sólo con *T* y, *G*, con *C*. Estas normas restrictivas, también llamadas de complementaridad, constituyen la base de la unión de las dos cadenas de ADN en una hélice bicatenaria, como la que forma el cromosoma que en *E. coli* porta la información genética.

El cromosoma de *E. coli* es un círculo de ADN bicatenario de unos cuatro

millones de nucleótidos de longitud. La secuencia de nucleótidos a lo largo de este círculo contiene, en clave, la información para fabricar varios tipos de ARN, así como todas las proteínas que se requieran en el metabolismo de la bacteria. La información puede estar contenida en cualquiera de las cadenas de la hélice de ADN. El segmento de nucleótidos que contiene la información para sintetizar un solo producto, ya sea proteína o una molécula de ARN ribosómico, se denomina gen. Los genes de los 55 componentes del ribosoma están desperdigados por el cromosoma, y representan alrededor del cinco por ciento de toda la información contenida en el ADN de *E. coli*.

La secuencia de nucleótidos que denominamos gen se transcribe en ARN con el concurso de un enzima llamado polimerasa de ARN. La polimerasa se une al ARN en un sitio específico y avanza a lo largo de una de las cadenas, "leyendo" la secuencia y catalizando la adición ordenada de nucleótidos a la cadena de ARN que se está sintetizando. Al igual que el ADN, el ARN es una secuencia de cuatro nucleótidos. Las subunidades que componen el



1. GENES DE LOS COMPONENTES RIBOSOMICOS, representados sobre un mapa cromosómico de la bacteria *Escherichia coli*. El cromosoma bacteriano es un anillo de ADN bicatenario formado por más de cuatro millones de bases nucleotídicas, sus subunidades. El círculo intermedio (*negro*) representa un mapa donde cada división corresponde a 40.000 bases. Se llama gen ribosómico al segmento de nucleótidos que lleva la información para fabricar un componente del ribosoma: una molécula de proteína o una molécula de ARN ribosómico (ARNr). El círculo externo cartografía los genes de las 52 proteínas del ribosoma. Las proteínas indicadas con una *S* pertenecen a la subunidad pequeña del ribosoma; las indicadas con *L*, a la subunidad grande. Entremezclados aparecen algunos genes no ribosómicos. En el punto *OriC* se inicia la replicación del cromosoma bacteriano. Los genes representados en una unidad monocolor se transcriben juntos en una molécula del ARN mensajero (ARNm), que posteriormente se traduce en proteína. Los genes para las tres moléculas de ARN ribosómico se representan en el círculo interno. Hay siete copias de cada gen de ARNr. Los 21 genes se disponen en siete juegos de tres. Para regular la síntesis del ribosoma debe ajustarse la producción de cada uno de los componentes en relación con la de los restantes 54.

ARN son las mismas que en el ADN, salvo que el uracilo (*U*) ocupa el lugar de la timina; como la timina, el uracilo es complementario de la adenina.

Los nucleótidos que la polimerasa va añadiendo al ARN son complementarios de los presentes en la cadena de ADN que se está leyendo. El ARN es, por tanto, una réplica en "negativo" de un segmento del cromosoma bacteriano y, como tal, conserva la información genética contenida en el gen. Una vez terminada la transcripción, la molécula de ARN recién sintetizada se separa del ADN.

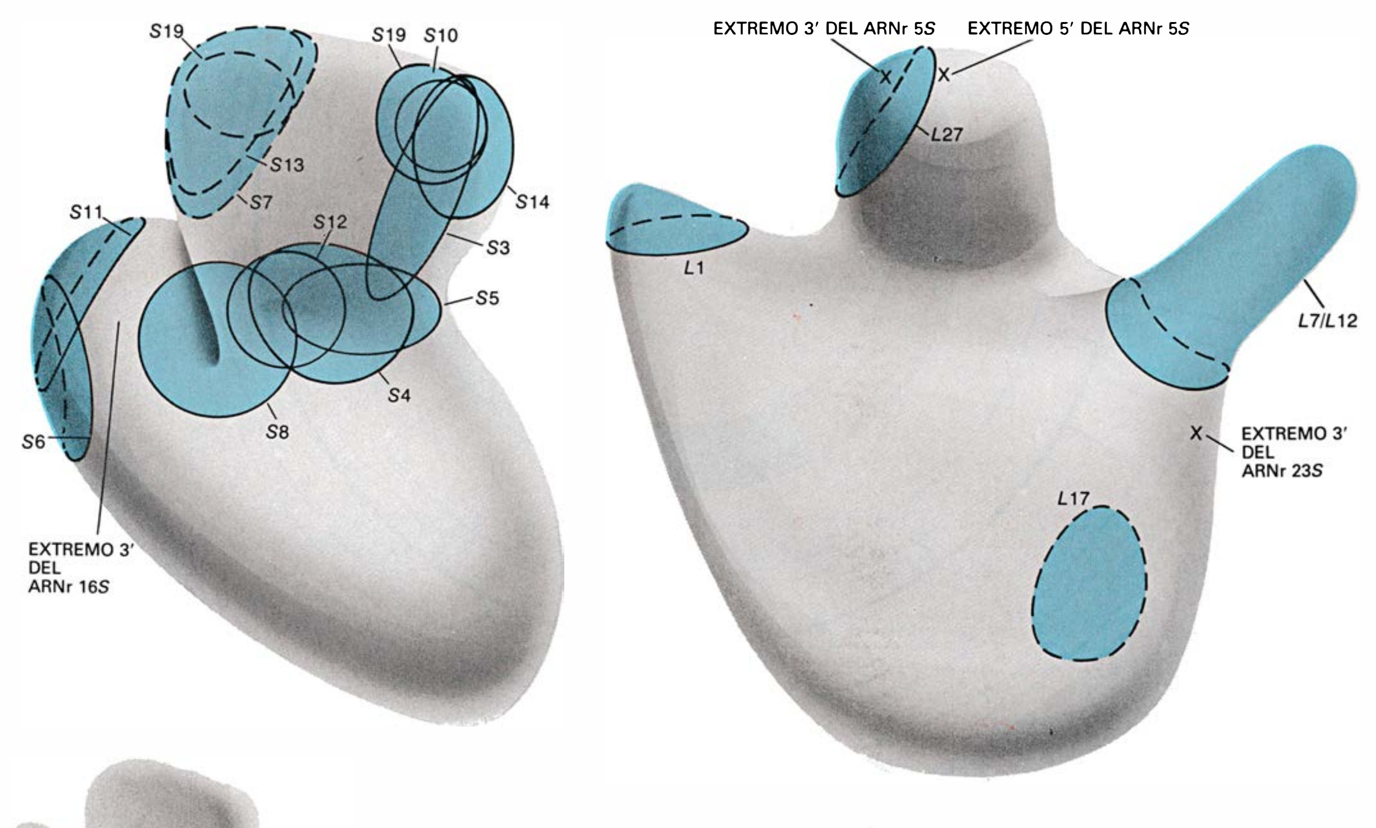
Si el ARN transcrito es ARN mensajero, los ribosomas se unen inmediatamente a él para traducir su mensaje. Ciertas secuencias breves del mensajero se utilizan para indicar el inicio y terminación de la traducción. Tras la unión del ribosoma al ARNm, el mensaje se lee de tres en tres nucleótidos. Cada "codón", o grupo de tres nucleótidos sucesivos, determina un aminoácido en la cadena de proteína.

Una pequeña molécula de ARN, el ARN de transferencia (ARNt), sirve de gancho para el empalme de aminoácidos a la proteína. Cada molécula de ARNt es específica para un aminoácido. La molécula posee, además, una secuencia de tres bases ("anticodón"), que marca la especificidad para el aminoácido, y que es complementaria a alguno de los codones presentes en el ARNm. Cuando el ARNt, por medio de dicha secuencia, reconoce a uno de los codones del ARNm, se une a él, en base a la relación de complementaridad, y transfiere el aminoácido al extremo de la cadena de proteína que se está sintetizando. Posteriormente, el ribosoma avanza hacia el siguiente codón, se expulsa el ARNt y se remplaza por uno nuevo, que lleva el aminoácido que se ha de añadir a continuación.

Como otras células bacterianas, *E. coli* es procariota, es decir, carece de núcleo diferenciado. En una célula procariota, transcripción y traducción están estrechamente ligadas. Muchas veces, los ribosomas empiezan a traducir ARNm antes de que éste se haya separado del ADN, continuando el proceso una vez que el ARNm se ha separado del todo. Por lo general, el ARNm tiene una vida media muy breve, degradándose enzimáticamente a los pocos minutos de haberse transcrito. Antes de que ello ocurra, el mensaje suele traducirse de 10 a 20 veces, de tal manera que de cada ARNm se puede sintetizar más de una molécula de proteína; el número exacto dependerá de la vida media del mensajero y de la velocidad de traducción.

Las células de organismos multicelulares, como nosotros, son eucariotas: poseen un núcleo definido y separado del citoplasma circundante por una membrana nuclear. En estas células, la transcripción se desarrolla en el núcleo; el mensajero se transporta posteriormente al citoplasma, donde se traduce. En las células eucariotas los ribosomas se ensamblan en el nucleolo, un orgánulo localizado en el interior del núcleo.

¿Cómo participan los procesos gené-



2. SUBUNIDADES DEL RIBOSOMA y localización de sus componentes. Cada subunidad consta de proteínas y ARNr. La subunidad pequeña contiene 21 moléculas de proteína, con una sola copia de cada una. La subunidad grande contiene 34 moléculas de proteína; una de ellas está presente con cuatro copias. Las regiones coloreadas indican la localización de algunas proteínas que exponen parte de su molécula en la superficie de la subunidad. Este mapa se dibujó a partir de los resultados obtenidos por inmunoelectromicroscopía; las zonas coloreadas corresponden a los sitios de la superficie del ribosoma que reconocen los anticuerpos específicos contra cada una de las proteínas. Los primeros datos en este sentido los aportaron el autor y sus colegas Lawrence Kahan y William A. Strycharz con James A. Lake, de la Universidad de California en Los Angeles; el trabajo lo continuaron Kahan y Lake. Georg Stoffler y colaboradores, del Instituto Max Planck de Genética Molecular, de Berlín, realizaron estudios similares. Hay una molécula de ARNr en la subunidad pequeña y dos en la grande. Las *X* indican el lugar de la superficie por donde aparecen los extremos de los ARNr. A la izquierda, el ribosoma funcional.

ticos fundamentales en la construcción de un ribosoma de una célula bacteriana? Los ribosomas mejor estudiados, con diferencia, son los de *E. coli*. Cada ribosoma de *E. coli* consta de dos subunidades de distinto tamaño y forma: la subunidad 30*S* y la 50*S*. (*S* significa unidades Svedberg, una medida de la movilidad de una partícula cuando se centrifuga a alta velocidad. El valor así definido es el coeficiente de sedimentación, una medida del tamaño de la partícula.)

La subunidad 30*S*, la más pequeña, contiene una molécula de ARN, el ARN ribosómico 16*S*, con una longitud de 1542 nucleótidos. Los otros componentes son proteínas, de las que hay 21 moléculas diferentes. Las proteínas se denominan *S*1 a *S*21; en este caso *S* no se refiere a la unidad Svedberg sino a subunidad pequeña (*small* en inglés). La molécula de ARN y las 21 proteínas están íntimamente unidas en una partícula de forma compleja y peso molecular de 900.000 dalton.

La subunidad 50*S*, la mayor, contiene dos ARN ribosómicos: una molécula 23*S* y otra 5*S*. Sus longitudes son 2904 y 120 nucleótidos, respectivamente. El resto de la subunidad está formado por 34 moléculas de proteínas. Las proteínas se denominan *L*1 a *L*34 (*L* por subunidad grande, *large* en inglés). Aunque en la subunidad 50*S* hay 34 moléculas de proteína, se agrupan en 31 tipos diferentes. De una de las proteínas hay cuatro copias, del resto una. No obstante, en el sistema de numeración se incluyen las 34, debido a errores iniciales de asignación que todavía se arrastran. Las proteínas y los dos ARNr conforman un esferoide aplanado, con varias protuberancias irregulares en su parte superior. La subunidad mayor tiene un peso molecular de 1,6 millones de dalton.

Las dos subunidades, tras su ensamblaje, forman el ribosoma funcional, que posee un coeficiente de sedimentación de 70*S* y un peso molecular de 2,5 millones de dalton. (Los coeficientes de sedimentación no son aditivos, ya que la forma de la partícula influye en su velocidad de sedimentación durante la centrifugación.) Una vez juntas, entre las subunidades grande y pequeña se forma un surco, que sirve de cobijo a enzimas y ARN mensajero durante la traducción.

Algunas de las primeras observaciones cuantitativas relacionadas con los ribosomas se referían a la cantidad de ARN presente en las células a diferentes tasas de crecimiento y división. Así, a finales de los años cincuenta, Ole Maaloe, Niels O. Kjeldgaard y sus colegas, del Instituto Serológico Estatal, y posteriormente en el Instituto Universitario de Microbiología de Copenhague, examinaron el contenido celular de la bacteria *Salmonella typhimurium*. Cultivaron *Salmonella*, una bacteria muy parecida a *E. coli*, en varios medios y encontraron una relación lineal entre la tasa de división de la bacteria y su contenido total de ARN. Por entonces se sabía ya que la mayoría del ARN celular era ARN ribosómico. La cantidad total de ARN es, pues, una medida aproximada de la concentración de ribosomas en la bacteria.

Trabajos posteriores confirmaron la existencia de una relación lineal entre la tasa de crecimiento y la cantidad de ribosomas de la célula: cuanto más rápidamente se dividen las bacterias, más ribosomas hay. (La relación lineal se pierde, no obstante, a tasas bajas de crecimiento.) La existencia de una relación lineal sugirió a los investigadores que el control de la síntesis de ribosomas debía ser crucial en la regulación del crecimiento y división celular. Algunos fisiólogos bacterianos pensaban que la tasa de crecimiento debía estar limitada en última instancia por la propia capacidad bacteriana para fabricar ribosomas; ello no es así, como veremos más adelante. No obstante, los trabajos del grupo de Copenhague sirvieron para centrar el interés en la regulación de la síntesis de los ribosomas.

Durante los años sesenta y principios de los setenta, nuestros trabajos se concentraron en el estudio de la estructura y función de los ribosomas y del ensamblaje del orgánulo a partir de sus numerosos componentes. A medida que profundizábamos, aumentaba nuestro interés por el problema, cada vez más complejo, de la regulación de la síntesis de los componentes ribosómicos. Cuando se aborda el problema del control surgen inmediatamente varias preguntas; por ejemplo, ¿está equilibrada y coordinada la síntesis de las diferentes partes del ribosoma?

Por equilibrada se entiende que los componentes se sintetizan en las mismas proporciones existentes en el producto final. Sea el proceso de fabricación de un producto industrial con dos componentes. Si necesitamos 50 unidades de dicho producto, una forma de asegurar esa cantidad será fabricar 50 unidades de cada una de las partes y ensamblarlas posteriormente. Suponiendo que no se rompa ninguna durante el proceso, una vez elaboradas las 50 unidades del producto el proceso de ensamblaje se parará por falta de ambos componentes. A eso llamamos control equilibrado.

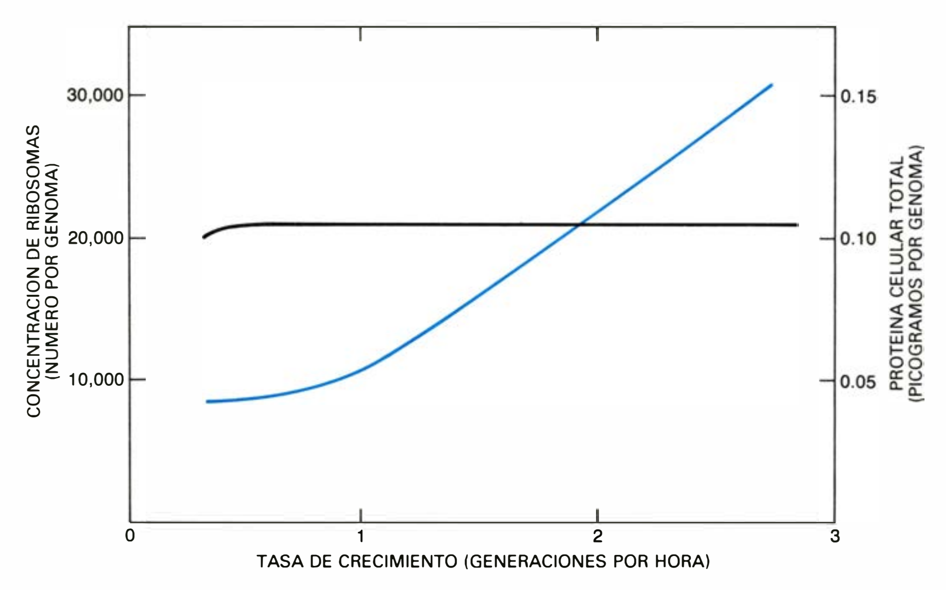
Regulación coordinada, por otra parte, significa que la síntesis de los componentes esté ajustada de forma que la relación entre las diferentes partes permanezca constante, con independencia de que el volumen total de producción aumente o disminuya. En nuestro ejemplo anterior, eso equivaldría a que, si se necesitasen 100 unidades del producto en vez de 50, habría que conseguir fabricar 100 unidades de cada uno de los componentes.

Podemos preguntarnos ahora si la producción de las 52 proteínas ribosómicas está equilibrada y coordinada y si se sintetizan en la misma proporción existente en el ribosoma. También cabe inquirir si la síntesis de tales proteínas está equilibrada y coordinada con respecto a la síntesis de los ARN ribosómicos, de suerte que todos los componentes del ribosoma se sinteticen siempre en cantidades similares. En caso contrario, la forma de regulación podría ser desequilibrada, existiendo excedentes de algunos componentes que serían posteriormente degradados por enzimas.

Estas fueron las primeras cuestiones que se plantearon al considerar en detalle la regulación de la síntesis de ribosomas. En las primeras etapas, el trabajo se centró en las proteínas ribosómicas. E. Kaltschmidt y Heinz-Günter Wittmann, del Instituto Max-Planck de Genética Molecular de Berlín, desarrollaron a la sazón un sencillo método para separar la mayoría de las proteínas ribosómicas. Este método y otros, elaborados posteriormente, fueron de gran utilidad en el estudio de la síntesis de las distintas proteínas ribosómicas a lo largo de diferentes fases del crecimiento.

En 1974, Patrick P. Dennis, becario postdoctoral en mi laboratorio de la Universidad de Wisconsin en Madison, observó, utilizando los nuevos métodos de separación de proteínas, que las tasas de síntesis de la mayoría de las proteínas ribosómicas estaban equilibradas, y que respondían de forma coordinada a los cambios de la composición del medio de cultivo. Ello bastaba para abordar directamente el problema de la regulación.

Desde los inicios de la investigación del ribosoma juzgué conveniente que, para estudiar la regulación de la síntesis de ribosomas en términos moleculares, se identificaran primero los genes de las proteínas y ARN ribosómicos. Una vez



3. CONCENTRACION DE RIBOSOMAS en *E. coli* (*línea de color*) en función de la tasa de crecimiento y la división celular; la relación es lineal. Para crecer y dividirse más rápidamente, la bacteria debe acelerar la síntesis de proteínas. Puesto que cada ribosoma tiene una capacidad limitada de síntesis de proteínas, *E. coli* debe fabricar más ribosomas cuando, por algún cambio del medio ambiente, haya de acelerar su crecimiento. Pero no puede despilfarrar ribosomas, sino que tendrá que ser capaz de ajustar con precisión sus propias necesidades. La concentración de proteínas (*línea negra*) no se incrementa al aumentar la tasa de crecimiento. Tanto la concentración de ribosomas como la de proteínas se expresan en relación a la masa del genoma bacteriano (su dotación génica). Son superiores en las células con crecimiento rápido, pero la relación proteína-ADN es igual que en las células con crecimiento lento.

identificado y aislado el ADN correspondiente a esos genes, podrían diseñarse experimentos de síntesis *in vitro* de proteínas y ARNr utilizando sistemas libres de células. Con este tipo de sistemas, y aplicando técnicas bioquímicas, podría medirse la tasa de expresión de los genes ribosómicos y dilucidar el mecanismo de control. Sin embargo, el aislamiento de los genes de los 55 componentes ribosómicos constituyó una tarea difícil que nos llevó mucho tiempo. No se disponía aún de los métodos de recombinación *in vitro* de ADN que permiten manipular material genético, y había que recurrir a la metodología de la genética clásica.

El bacteriófago lambda, un virus que infecta bacterias, fue clave en las investigaciones. Después de invadir la célula bacteriana, el fago inserta su ADN en el cromosoma del hospedador. Tras la inserción puede coexistir pacíficamente con la bacteria durante mucho tiempo. Sin embargo, en ciertas condiciones, por ejemplo ante daños causados al ADN del hospedador por radiaciones ultravioleta, el virus puede abandonar el cromosoma bacteriano y autorreproducirse.

El primer paso en la reproducción del fago es su escisión del cromosoma bacteriano. En algunos casos, la escisión es incorrecta y junto con el ADN del fago puede ir también un segmento del cromosoma bacteriano. El ADN escindido se introduce en las nuevas partículas víricas. Si junto con el ADN vírico se introduce ADN bacteriano, el fago puede servir para "transducir" ADN bacteriano a otra bacteria. Cuando empezaron los trabajos para aislar los genes ribosómicos se tenía conocimiento de que muchos de los genes de las proteínas ribosómicas se agrupaban en un segmento del cromosoma de *E. coli*. Tal región constituía un blanco excelente para un fago transductor.

A finales de 1974, S. Richard Jaskunas, Lasse A. Lindahl, un becario postdoctoral de mi laboratorio, y yo conseguimos que el fago se integrase en el ADN bacteriano, cerca de los genes de las proteínas ribosómicas. A continuación aislamos varias partículas de fagos que llevaban uno o más genes de proteínas ribosómicas. El número variaba considerablemente: aislamos un fago que llevaba 27 genes, más de la mitad del total.

Este fue el primer paso importante en los estudios sobre la regulación de la síntesis de ribosomas. Su trascendencia residía en que permitía la construcción de un sistema *in vitro* donde estudiar cómo se regulaba la síntesis de los componentes ribosómicos. Posteriormente, otros investigadores, algunos de mi laboratorio entre ellos, aislaron genes de proteínas ribosómicas presentes en otras regiones del cromosoma de *E. coli*.

Simultáneamente al estudio de los genes de proteínas ribosómicas, a mediados de los setenta, se inició el de los genes para los ARNr, fundamentalmente por parte de Norman R. Davidson y sus colaboradores, del Instituto de Tecnología de California; otros laboratorios, el nuestro incluido, también participaron en la tarea. A diferencia de los genes de proteínas ribosómicas, de los que hay una copia de cada uno en el cromosoma de *E. coli*, existen siete copias de cada uno de los tres genes para ARNr. Los genes de cada juego se transcriben juntos en un único ARN y, posteriormente, tras un proceso de maduración, resultan las moléculas 16*S*, 23*S* y 5*S*. En 1977 se habían aislado ya los siete juegos de genes de ARNr, y se trabajaba en la determinación de su estructura y secuencias nucleotídicas.

Una vez purificados los genes de los ARNr y gran parte de los de proteínas, podíamos abordar el problema de su regulación. Queríamos saber cómo se regulaba un conjunto de 55 genes para conseguir una síntesis equilibrada y coordinada de todos los componentes ribosómicos. En aquellos tiempos el concepto de regulación genética con mayor aceptación era la teoría del operón, propuesta a principios de los años 60 por Jacques Monod y François Jacob, del Instituto Pasteur de París, para explicar cómo las células de *E. coli* controlaban el metabolismo de la lactosa.

E. coli puede aprovechar varias moléculas como fuente de carbono; entre ellas, los azúcares lactosa y glucosa. Para utilizar lactosa ha de disponer del enzima betagalactosidasa. Si las bacterias se cultivan en un medio donde la lactosa constituye la fuente de carbono, se acumula betagalactosidasa en la célula. La utilización de glucosa, sin embargo, no requiere betagalactosidasa, y si las bacterias se cultivan en un medio con glucosa como fuente de carbono el nivel de betagalactosidasa es muy bajo.

Monod y Jacob demostraron que había tres genes implicados en el metabolismo de la lactosa, uno de los cuales era el gen de la betagalactosidasa. Los genes forman una unidad genética: el operón de la lactosa (operón lac). Los genes, que están contiguos en el cromosoma, se transcriben juntos. La transcripción del operón lac comienza cuando un enzima polimerasa se une a un segmento de nucleótidos que recibe el nombre de promotor lac, cercano a uno de los extremos del operón.

Junto al promotor hay un grupo de nucleótidos denominado operador. Una proteína, el represor lac, puede unirse al operador. Cuando el represor

está unido al operador, la polimerasa no puede hacerlo con el promotor y no se transcribe el operón lac. Esta es la situación cuando no hay lactosa en la célula bacteriana.

Sin embargo, cuando entran en la bacteria moléculas de lactosa, sufren una rápida modificación. La forma modificada, estructuralmente muy parecida a la lactosa, actúa de inductor: se une al represor lac e impide que el represor reconozca al operador. El promotor puede entonces interaccionar libremente con la polimerasa y se inicia la transcripción. La presencia de lactosa constituye, por tanto, la señal que activa el operón; su ausencia, la señal que lo reprime.

El éxito del modelo del operón para explicar la regulación del metabolismo de la lactosa centró la atención de la mayoría de investigadores en el control directo de la transcripción. En parte por ello, a finales de los años 60 y principios de los 70, la mayoría de los sistemas reguladores conocidos se basaban en un control de la transcripción mediante mecanismos similares al del modelo original del operón. También en nuestro caso se intentó aplicar la teoría del operón.

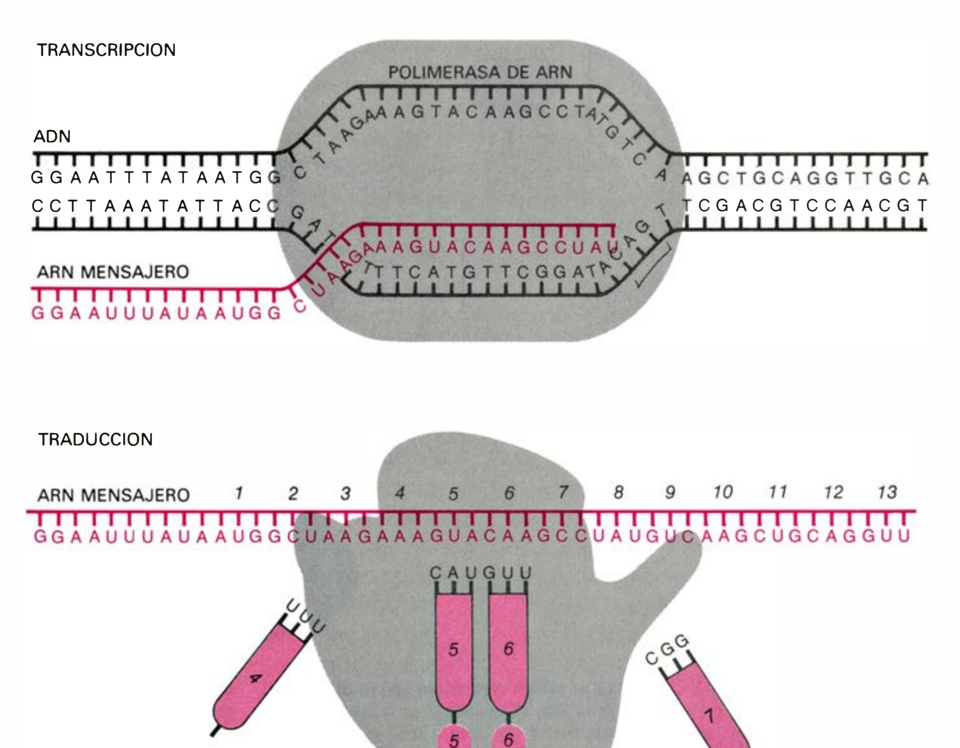
Para comprobar si dicha teoría se acomodaba a la síntesis de ribosomas había que identificar primero las unidades de transcripción y promotores de los componentes ribosómicos. Se sabía que los genes de los ARN ribosómicos estaban organizados en siete juegos, cada uno de los cuales constituía un operón, o unidad de transcripción. El trabajo de mi laboratorio se centró, por tanto, en los genes de las proteínas ribosómicas. Pronto descubrimos que dentro de los principales agrupamientos de genes de proteínas los genes se organizaban en operones, cada uno de los cuales contenía uno o varios genes que se transcribían juntos a partir de un único promotor, y que cada operón podía transcribirse independientemente.

Conocida la disposición funcional de los genes de proteínas ribosómicas, podía avanzarse una hipótesis sobre su regulación. La hipótesis, en la línea del modelo del operón, se basaba en el control directo de la transcripción. Aspecto central de la hipótesis era la suposición de que todos los operones debían responder a las mismas señales químicas (moléculas inductoras y represoras). La presencia del inductor activaría todos los operones simultáneamente y su ausencia los reprimiría a todos.

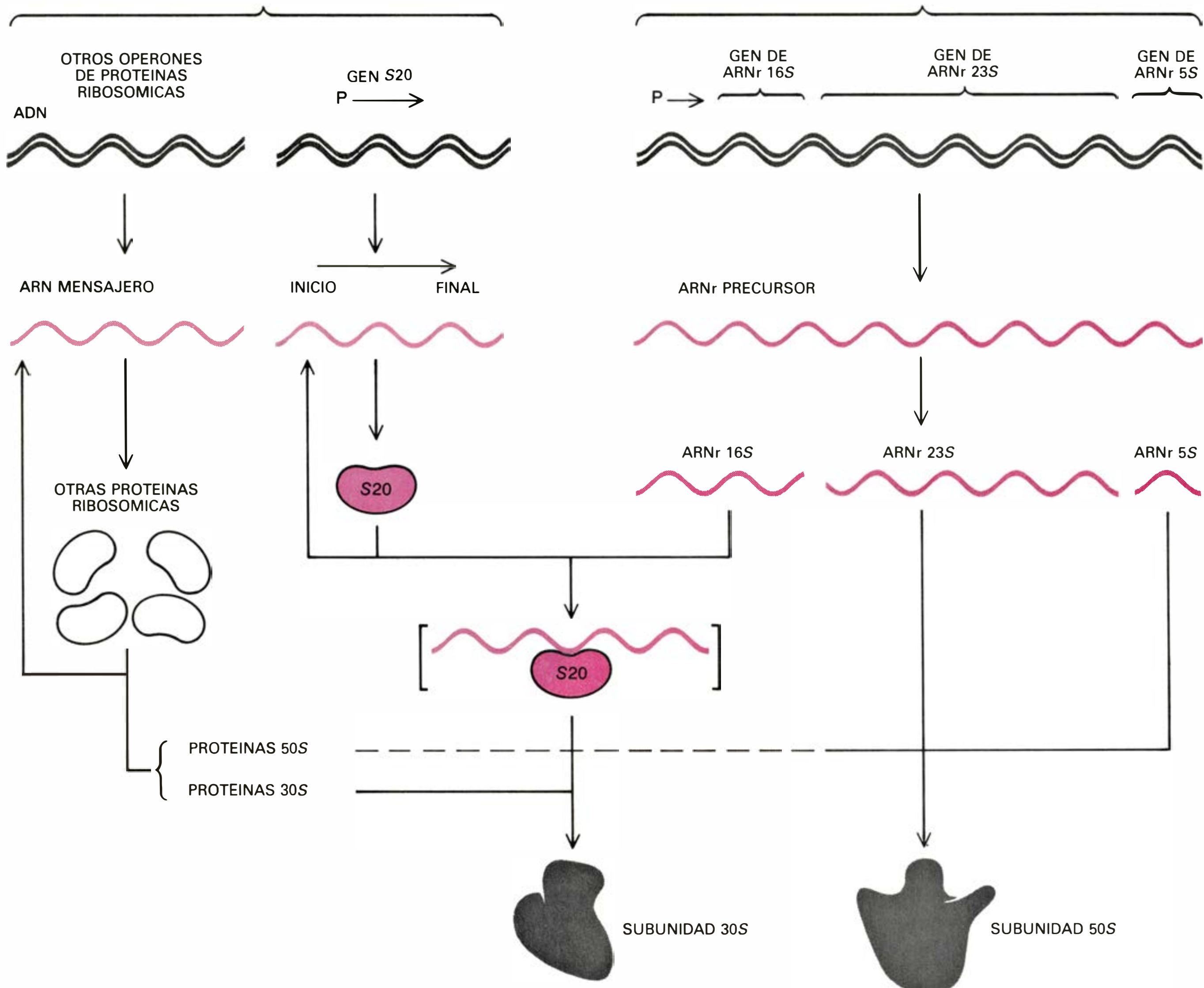
De ser cierto que todos los operones de proteínas ribosómicas responden a las mismas señales, las secuencias nucleotídicas de los promotores y operadores habrían de presentar ciertas características comunes. La similitud de los promotores y la simultaneidad de su respuesta a un único inductor explicarían la síntesis coordinada de las proteínas ribosómicas.

Aunque la teoría del operón se formulara para explicar la acción de unos pocos genes adyacentes, no había razón, en principio, para pensar que un mecanismo similar no controlara la acción de un conjunto nutrido de genes, incluso genes esparcidos por el cromosoma bacteriano. La verdad es que se encontraron ejemplos de sistemas complejos de operones, uno de ellos formado por genes que se activaban cuando el cromosoma de *E. coli* sufría algún daño.

No obstante, la regulación de la síntesis de ribosomas por un mecanismo de tipo operón debería ajustarse a ciertos requisitos. Para explicar la síntesis equilibrada de las proteínas ribosómicas era necesario admitir que todos los operones implicados se transcribían con la misma eficacia, que todas las moléculas de ARN mensajero fabrica-



4. TRANSCRIPCION Y TRADUCCION son los procesos genéticos fundamentales en la síntesis de un ribosoma. En la transcripción (*arriba*) se copia la información genética del ADN en ARN. Los cuatro nucleótidos del ADN son adenina, citosina, guanina y timina (*A, C, G* y *T*). En el ARN, uracilo (*U*) sustituye a timina. *A* puede formar enlaces sólo con *T* o *U* y, *C*, con *G*. Las bases que establecen enlaces se denominan complementarias. En la transcripción, el enzima polimerasa de ARN se une al cromosoma bacteriano. A su tiempo, va separando las cadenas de ADN. Se sirve de una cadena como molde para fabricar una cadena de ARN mensajero o ribosómico, a partir de nucleótidos libres. Por traducción (*abajo*) se entiende el proceso de unión de aminoácidos para formar una cadena de proteína, de acuerdo con la información que lleva el ARNm. Cada codón, o grupo de tres nucleótidos del ARNm, determina un aminoácido. Las pequeñas moléculas de ARN de transferencia (ARNt) portan anticodones: tres bases complementarias al codón. Cada ARNt se une específicamente a un aminoácido. El ribosoma avanza a lo largo del ARNm, leyendo los codones ordenadamente. Por el anticodón, el ARNt se une al codón, transfiriendo el aminoácido que lleva en uno de sus extremos a la cadena de proteína que se está formando. Cuando el ribosoma se desplaza al siguiente codón se expulsa el ARNt; otro nuevo ocupará su lugar.



5. RETRORREGULACION, sistema que utiliza la proteína *S*20 para controlar su propia síntesis. Los genes de proteínas ribosómicas se transcriben en grupos, como operones, unidades compuestas por uno o más genes que se transcriben juntos. Cada operón posee su propio promotor, o segmento de nucleótidos por donde comienza la transcripción. Una de esas unidades, el gen *S*20, se transcribe a partir de *P*, su promotor. Cada juego de tres genes de ARN ribosómico constituye un operón. Los ARNr se transcriben juntos en una molécula precursora que posteriormente se divide, produciendo las tres moléculas de ARNr. *S*20 es una proteína capaz de unirse específicamente a un sitio determinado; en el proceso de ensamblaje del ribosoma forma un complejo con el ARNr 16*S* y se incorpora finalmente a la subunidad 30*S*. Mientras haya ARNr 16*S* disponible se sintetizará proteína *S*20. Al agotarse el ARNr 16*S*, el exceso de moléculas *S*20 reprimirá la síntesis de más proteínas uniéndose a un sitio específico del ARN mensajero y bloqueando su traducción.

das por esos operones tendrían la misma vida media y que la tasa de traducción de todas las secuencias informativas de esos mensajeros sería la misma. De no cumplirse alguna de esas condiciones, las proteínas ribosómicas se sintetizarían en cantidades diferentes.

Se sabía, además, que la tasa de síntesis de ARNr y proteínas ribosómicas venía a ser la misma, al menos en células que crecían normalmente. Para sostener la hipótesis del operón había que suponer que los promotores de los operones de los ARNr compartían características comunes con los promotores de los operones de las proteínas ribosómicas. Era preciso que los operones de las proteínas respondiesen coordinadamente a las mismas señales. Además, los ARNr se transcriben directamente del ADN, mientras que las proteínas se sintetizan a partir de las moléculas de mensajeros, que pueden traducirse más de una vez, por lo que la eficacia de los promotores de los operadores de los ARNr habría de ser mayor que la de los promotores de los operones de proteínas, y la relación entre ellos debería ser tal que se fabricase la misma cantidad de los dos tipos de componentes.

Tamañas dificultades teóricas no parecían insalvables y, en cualquier caso, el modelo del operón constituyó un valioso punto de partida. Si, en efecto, hubiese un mecanismo de tipo operón, cabría esperar, como ya hemos dicho, similitudes entre los promotores de todos los operones ribosómicos. Afortunadamente, casi al mismo tiempo que se aislaban los genes y promotores ribosómicos, se pusieron a punto dos nuevas técnicas para determinar secuencias de nucleótidos, una de ellas desarrollada por Frederick Sanger y sus colegas, del Laboratorio de Biología Molecular del Consejo de Investigaciones Médicas, en Cambridge, y la otra por Allan M. Maxam y Walter Gilbert, de la Universidad de Harvard.

Leonard E. Post, alumno de mi laboratorio, determinó con las nuevas técnicas la secuencia completa de cinco de

los promotores de los operones de las proteínas de los ribosomas. También en mi laboratorio, Herman A. deBoer y Scott F. Gilbert determinaron la secuencia de nucleótidos de los promotores de dos de los juegos de genes de ARN ribosómicos. Otros investigadores hicieron lo mismo con los restantes juegos de genes de ARNr.

El trabajo de secuenciación no descubrió ninguna característica común notable que distinguiera a los promotores ribosómicos de otras secuencias de promotores conocidos. La hipótesis de un mecanismo de tipo operón para explicar la síntesis equilibrada y coordinada de los componentes ribosómicos empezaba a perder fuerza.

Puesta en duda la regulación directa del conjunto de genes ribosómicos, quedaban dos posibilidades alternativas. De acuerdo con la primera, la síntesis de ARN ribosómico estaría regulada directamente y la propia síntesis de los ARNr controlaría, de forma secundaria, la síntesis de proteínas. La segunda posibilidad era: los genes de proteínas ribosómicas se regularían directamente; esta regulación condicionaría, de forma secundaria, la de los genes de ARNr.

La segunda posibilidad presentaba los mismos inconvenientes que el modelo básico del operón, ya que son muchos los operones de proteínas ribosómicas. Implicaba, por tanto, la existencia de muchos promotores con la misma estructura, posibilidad que ya habíamos descartado. El primero de los mecanismos, sin embargo, parecía factible. Una forma sencilla de controlar la tasa total de síntesis de ribosomas era sintetizar cantidades ligeramente mayores de proteínas que de ARN. El mismo ARN constituiría un factor limitante a la hora de ensamblar ribosomas. El pequeño exceso de proteínas de ribosoma actuaría de represor, interrumpiendo la síntesis de más proteínas. Se admitía que la regulación de la síntesis de proteínas ribosómicas se llevaba a cabo a nivel de la transcripción.

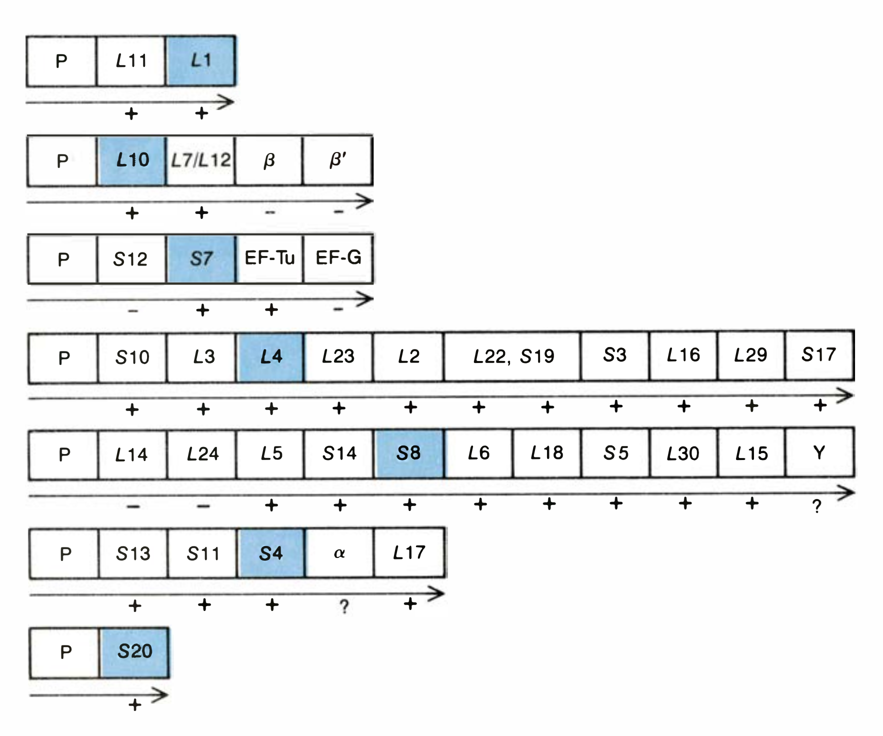
Ann M. Fallon, Sue Jinks-Robertson y Geneva D. Strycharz, colaboradores míos, realizaron los primeros experimentos que trataban de averiguar si la síntesis de proteínas de ribosoma estaba controlada por un mecanismo de retroinhibición. Además, se pretendía comprobar definitivamente la validez de la hipótesis del operón. Si el fundamento de la síntesis equilibrada de las proteínas ribosómicas fuera la existencia de promotores idénticos, sensibles al mismo inductor o represor, la inclusión, en la bacteria, de más copias de algunos de los operones de proteínas, con su correspondiente promotor, desequilibraría el proceso de síntesis. Las proteínas correspondientes a los genes añadidos se sintetizarían en mayor cantidad que el resto.

Sin embargo, si la síntesis fuese equilibrada y estuviese condicionada por su mecanismo de retroinhibición, el efecto de incremento de "dosis génica" no se apreciaría, ya que el mecanismo de retroinhibición reprimiría la acción de los genes añadidos, de la misma forma que lo haría con los genes de la propia bacteria. Es decir, se mantendría equilibrada la síntesis.

Se introdujeron en *E. coli* copias de algunos genes de proteínas de ribosoma, utilizando como vehículos el bacteriófago lambda o un plásmido (pequeña molécula autónoma de ADN circular). Como hemos visto, el bacteriófago puede infectar la célula bacteriana y coexistir en ella. Los plásmidos también pueden penetrar en la bacteria. Las copias adicionales de genes de proteínas ribosómicas presentes en el fago o en el plásmido se expresan junto con la información presente en el cromosoma.

Una vez introducidos en las bacterias el fago o los plásmidos con los genes extra, se medía la tasa de transcripción y de traducción. La tasa de transcripción de los ARN mensajeros para proteínas se incrementaba en proporción al número de copias de genes presentes. La tasa de traducción, sin embargo, no se alteraba. Las proteínas cuyos genes se habían introducido con el fago o los plásmidos no se sintetizaban en mayor cantidad que las proteínas cuyos genes tenían un número normal de copias.

Estos resultados se ajustaban a la idea de regulación por retroinhibición. Sin embargo, resultaba evidente que el control se estaba ejerciendo sobre la traducción, no sobre la transcripción. Se conocían algunos casos de control genético basados en la regulación de la traducción, lo que no restaba sorpresa a nuestros descubrimientos, inesperados, pues estábamos acostumbrados a la noción de control transcripcional. Después de repetir los experimentos muchas veces, propusimos el modelo de retroinhibición de la traducción para explicar la síntesis equilibrada de las proteínas del ribosoma. Casi

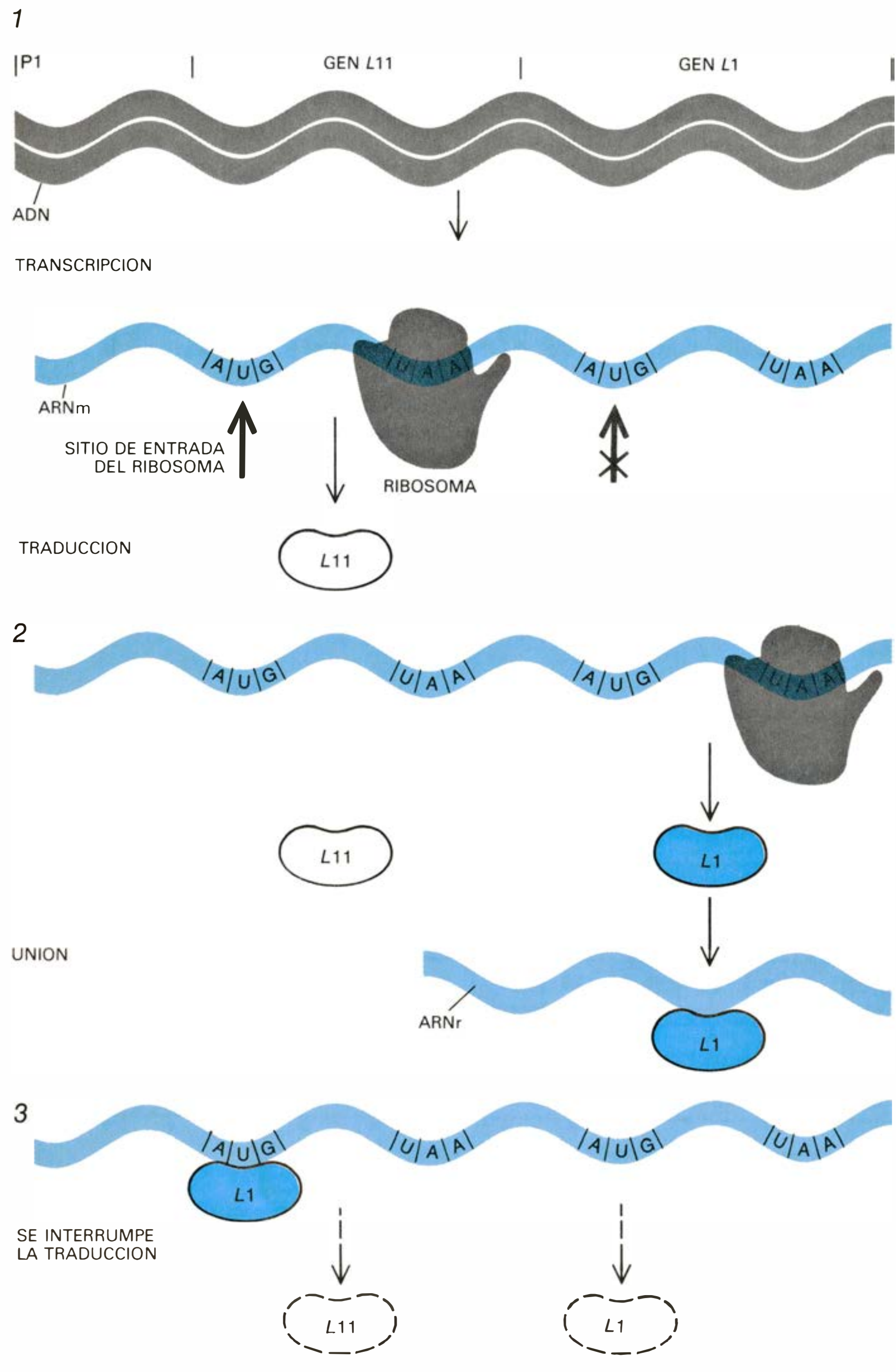


6. EL REPRESOR DE LA TRADUCCION es una proteína del ribosoma capaz de inhibir su propia síntesis y, a veces, la de otras proteínas ribosómicas. Lo logra uniéndose al ARNm de su propio operón, con lo que detiene su traducción. En los casos aquí indicados, los represores son las casillas coloreadas. Cada hilera representa un operón de proteínas ribosómicas. La transcripción empieza en el promotor (*P*); las flechas indican el sentido de la transcripción. En cada operón, las casillas marcadas con un signo + corresponden a proteínas reprimidas por el represor. No todas las proteínas de un operón están sujetas a ese tipo de represión. Por ejemplo, *S*7 no inhibe la síntesis de *S*12. Parece que *S*12, *L*14 y *L*24 pueden interrumpir su propia síntesis. El operón se divide, por tanto, en unidades reguladoras. Los operones ribosómicos pueden incluir genes para proteínas que no forman parte del ribosoma. Alfa, beta y beta' son subunidades de la polimerasa de ARN. *EF-Tu* y *EF-G* son "factores de elongación", necesarios para la síntesis de proteínas. (No deben confundirse estos operones con los que afectan a la transcripción génica.)

al mismo tiempo, Dieter Geyl y August Böck, de la Universidad de Regensburg, y Niels Fiil, del Instituto Universitario de Microbiología de Copenhague, en colaboración con Patrick Dennis, que ahora trabaja en la Universidad de la Columbia Británica, realizaron experimentos similares. Sus resultados apoyaban la teoría que propugnaba un control de la traducción por retroinhibición.

Según dicha teoría, las proteínas ribosómicas libres (no unidas a ARNr) inhiben la traducción de sus propios mensajeros. La síntesis de proteínas de ribosoma está acoplada al proceso de ensamblaje de ribosomas. Así, mientras se precisan proteínas para construir ribosomas (es decir, mientras haya ARNr al que unirse), los mensajeros se traducirán continuamente. Cuando no quede ARN al que engarzarse, las proteínas libres reprimirán la traducción de sus propios mensajeros.

Las pruebas directas en apoyo de esta teoría se recabaron en los experimentos *in vitro* realizados en mi laboratorio por John L. Yates, quien mezcló en un tubo de ensayo ADN con información para sintetizar proteínas ribosómicas, ribosomas, polimerasas de ARN y otros enzimas y moléculas necesarios para la transcripción y traducción. Además, añadió aminoácidos marcados con átomos radiactivos, al objeto de detectar las proteínas sintetizadas a partir de dichos aminoácidos.



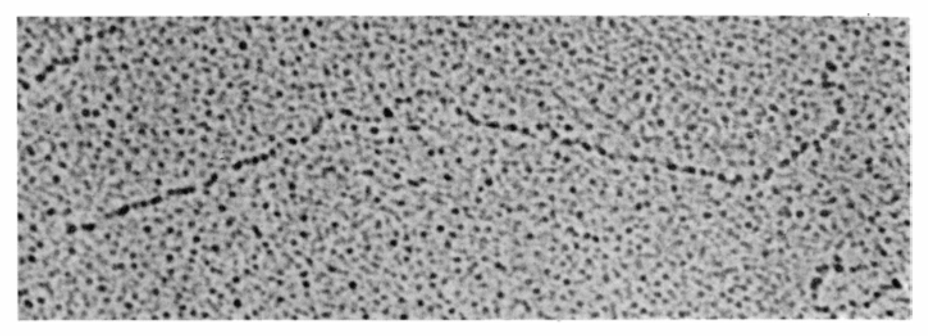
7. TRADUCCION SECUENCIAL que, junto con la retroinhibición de la traducción, asegura que las proteínas ribosómicas se sinteticen en cantidades iguales. El operón *L*11-*L*1 contiene un promotor y los genes de las proteínas *L*11 y *L*1, en este orden (*1*). La secuencia *AUG*, situada al principio de cada secuencia informativa del ARN mensajero, sirve de señal para incoar la traducción; *UAA*, al final de la secuencia, es la señal para detenerla. El ribosoma sólo puede iniciar la traducción del mensaje *L*11-*L*1 empezando por *L*11, en el sitio denominado punto de entrada del ribosoma. Sólo después de haber traducido *L*11 puede traducir *L*1 (*2*). Una vez sintetizadas las dos proteínas, *L*1 se une al ARN ribosómico 23*S* en tanto haya copias de éste. Cuando el ARNr 23*S* se agota, *L*1, un represor de la traducción, se une al ARNm *L*11-*L*1 cerca del sitio de entrada del ribosoma y detiene la traducción (*3*). De esta forma, *L*11 y *L*1 se sintetizan siempre en igual cantidad. El bloqueo de la traducción anula la síntesis de ambos.

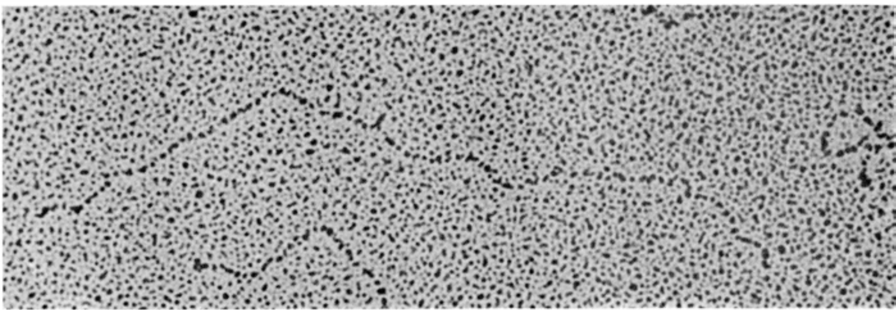
En 1979 se había purificado la mayoría de las proteínas ribosómicas. Cuando a la mezcla anteriormente descrita se añadieron pequeñas cantidades de proteínas purificadas, se observó con sorpresa que algunas proteínas inhibían la traducción de sus correspondientes ARN mensajeros. No todas las proteínas tenían esa propiedad, pero las que la ofrecían eran capaces de inhibir la síntesis de otras proteínas, además de la suya propia. Los genes de proteínas cuya síntesis se interrumpía estaban situados en el mismo operón que el gen de la proteína represora. (A este represor se le denomina represor de la traducción, para distinguirlo del represor de la transcripción, del clásico modelo del operón.) Por ejemplo, cuando se añadía al tubo de ensayo la proteína *S*4, no sólo se detenía su propia traducción, sino también la correspondiente a las proteínas *S*13 y *S*11; las tres proteínas están en el mismo operón. La proteína *L*1 inhibe su síntesis y la de *L*11, con quien comparte una unidad de transcripción. Sin embargo, la presencia de la proteína *L*1 no afecta a la síntesis de la proteína *S*4, *S*11 o *S*13; la presencia de *S*4 no tiene efecto sobre la traducción de *L*1 o *L*11.

Las proteínas *S*8, *S*4 y *L*1 se identificaron inmediatamente como represores de la traducción. Al poco tiempo, en mi laboratorio, identificamos también como tales a *S*7 y *L*4. Casi simultáneamente, y con independencia, Ryuji Fukuda, de la Universidad de Kyoto, y Herbert Weissbach y sus colegas, del Instituto Roche de Biología Molecular, demostraron que *L*10 podía detener la traducción de su propio mensajero. Böck y Reinhard Wirth, de Regensburg, demostraron que también *S*20 era un represor. Cada represor regula la

traducción de todas las proteínas de su operón o de un subgrupo de ellas. Cada uno de los grupos de proteínas reguladas de esta forma recibe el nombre de unidad de regulación, y hay un solo represor por cada unidad de regulación. Al descubrirse la represión de la traducción parecía que el represor debía detener la traducción uniéndose al ARN mensajero correspondiente, impidiendo así que el ribosoma se uniera al ARNm para comenzar a traducir.

Dado que cada proteína represora podía inhibir la traducción de varias proteínas ribosómicas, todas determinadas por un mismo mensajero, la represión podría lograrse de dos formas, bien uniéndose el represor a un único sitio del ARNm e impidiendo la traducción de todas las secuencias informativas de una vez, o bien interaccionando con varios sitios del ARNm, uno por cada secuencia informativa.





8. MOLECULAS DE ARN, en forma de largas moléculas monocatenarias, y en condiciones que impiden el emparejamiento entre regiones adyacentes. La micrografía superior muestra ARN ribosómico 16*S* en esas condiciones. Cuando se mezcla con la proteína *S*8, un represor de la traducción, se forma una pequeña zona bicatenaria, hacia la mitad de la cadena de ARN, como se muestra abajo. La proteína *S*8, que no se aprecia en la micrografía, está unida a la región bicatenaria. El mecanismo de retrorregulación se basa en la capacidad de *S*8 para formar complejos con el ARNr 16*S* (en la síntesis de ribosomas) o con el ARN mensajero de su operón (para bloquear su propia síntesis). Las electromicrografías son obra de Michael Beer, de la Universidad Johns Hopkins, Theo Koller, del Instituto Federal Suizo de Tecnología, y de investigadores adscritos al laboratorio del autor, en la Universidad de Wisconsin en Madison.

Las investigaciones realizadas en mi laboratorio por Yates, Dennis Dean y Gail A. Baughman sugieren que hay un único sitio de unión en cada unidad de regulación y que la unión del represor a ese sitio interrumpe simultáneamente la traducción de todas las proteínas de la unidad. Ello parece indicar que el ribosoma accede a la traducción del mensajero por un único sitio, el denominado sitio de entrada del ribosoma. Es necesario que se traduzca la primera secuencia informativa situada inmediatamente tras el sitio de entrada del ribosoma para que se traduzcan las que vengan a continuación.

Este fenómeno de traducción secuencial, o acoplada, se comprobó por primera vez en el operón formado por los genes *L*11 y *L*1. En el mensajero *L*11-*L*1, el sitio de entrada del ribosoma precede a la secuencia informativa de *L*11; la proteína *L*1 se sintetiza una vez terminada la *L*11. Probablemente el mismo ribosoma realice la síntesis de ambas proteínas de forma secuencial. Por cada molécula de *L*11 se traduce otra de *L*1.

El principio de traducción secuencial probablemente tenga validez general en los casos de síntesis de proteínas de ribosoma controlada por retroinhibición de la traducción. De hecho, se ha demostrado su existencia incluso en algunos operones no ribosómicos. En estos casos, la unidad de regulación está formada por varios genes y la traducción de las secuencias informativas presentes en el ARN mensajero se realiza siempre de manera secuencial.

Se aprecia inmediatamente que un proceso de traducción secuencial junto con el mecanismo de represión a nivel de traducción determinan una síntesis equilibrada y coordinada de proteínas ribosómicas en relación a los ARNr. La represión de la traducción asegura que las proteínas de ribosoma se sinteticen en la misma cantidad que los ARNr: cuando se agota el suministro de ARNr, las proteínas represoras se unen inmediatamente a su propio ARN mensajero y se detiene la síntesis de nuevas proteínas ribosómicas. Si posteriormente se incrementa la tasa de síntesis de ARNr, su disponibilidad para fabricar más ribosomas determina que empiece de nuevo la traducción de los ARNm de las proteínas ribosómicas. Una vez en marcha, el mecanismo de traducción secuencial asegura la síntesis de cantidades similares de todas las proteínas de ribosoma.

Uno de los aspectos del mecanismo de represión de la traducción que mayor interés suscitó fue la interacción entre represor y mensajero. La importancia de tal interacción radica en que, una vez iniciado el ensamblaje del ribosoma, se dispone tanto de ARNr como de los mensajeros para proteínas ribosómicas. Un punto esencial de la teoría de la retroinhibición de la traducción es que ARNr y ARNm deben competir por unirse a la proteína ribosómica represora. Para que exista tal competencia, la proteína represora debe ser capaz de unirse a cualquier molécula de ARN. Gracias a los estudios sobre ensamblaje de ribosomas se conocían bastante bien las interacciones entre proteínas y ARNr, pero se sabía poco, sin embargo, sobre cómo interaccionaban represor y mensajero.

El proceso se fue esclareciendo a medida que avanzaban las investigaciones y nuevos datos se añadían a los ya conocidos. Los primeros represores identificados en los trabajos sobre síntesis de ribosomas pertenecían todos a esa clase especial de proteínas cuya función es la de unirse a un sitio específico. Tales proteínas formaban, in vitro, complejos específicos con el ARN en ausencia de otras proteínas ribosómicas. Tal tipo de interacciones se conocía con anterioridad al descubrimiento de la represión de la traducción; además, se habían caracterizado algunos de los sitios que dichas proteínas utilizan para unirse a la molécula de ARN.

La información recogida sobre tales sitios fue muy útil para entender cómo compiten ARN ribosómicos y mensajeros por unirse a la proteína represora. El modelo más sencillo es que sea la misma región de la proteína represora

la que interaccione con cualquiera de los dos ARN. Si este fuera el caso cabría esperar ciertas similitudes estructurales entre los correspondientes sitios de unión situados en los ARNr y ARNm.

En los sitios de enlace de los ARNr y ARNm que interactuaban con el mismo represor, las similitudes eran de dos tipos. Por una parte, había secuencias nucleotídicas similares en ambos tipos de ARN. Además, las estructuras secundarias de las cadenas de ARN eran también semejantes. Se suele entender por estructura secundaria de una molécula de ARN la disposición que adopta en el espacio. La configuración que resulta al plegarse la cadena viene condicionada, en parte, por la posibilidad de que algunos segmentos de la cadena de ARN sean complementarios y formen estructuras bicatenarias. Tales estructuras pueden formar parte de sitios donde sólo una proteína represora específica constituye un complejo con el ARN.

Las similitudes entre los correspondientes sitios de los ARN mensajeros y ribosómicos permiten la actuación de la retroinhibición. Mientras haya ARNr transcribiéndose, las proteínas ribosómicas represoras se unirán a ellos. Cuando la tasa de transcripción de ARNr disminuya, el represor se acumulará ligeramente. Algunas moléculas del represor no podrán encontrar más ARNr al que unirse y reconocerán el sitio de unión presente en el ARNm, formando un complejo con él e interrumpiendo su traducción. Puesto que se sigue ese mismo proceso simultáneamente con todas las proteínas represoras, se detendrá la traducción de todas las proteínas ribosómicas.

Mientras haya ARN ribosómicos y mensajeros, la probabilidad de que el represor se una al ARNr debe ser considerablemente mayor de que se una al ARNm. En otras palabras, la unión del represor al ARNr debe estar, en términos energéticos, más favorecida que la unión del represor al ARNm. Si no fuera así, la unión del represor al mensajero interrumpiría la traducción, independientemente de que hubiera ARNr disponible para formar nuevos ribosomas. El sistema constituido por el represor y los dos ARN dejaría de funcionar como mecanismo de retroinhibición.

Hay experimentos que apoyan esta teoría. La acción represora de la proteína *L*1 sobre la síntesis de las proteínas *L*11 y *L*1 puede suprimirse in vitro añadiendo ARN ribosómico 23*S* a la mezcla reactiva. Se sabe que el ARNr 23*S* forma complejo con *L*1 durante el ensamblaje del ribosoma, y hemos visto que *L*1 reprime la expresión de su propio operón, impidiendo la traducción. De este experimento se desprende, por tanto, que la formación del complejo *L*1-23*S* está mucho más favorecida que la del conjunto entre *L*1 y el ARNm del operón *L*11-*L*1.

Los mecanismos de control retroactivo basados en una relación competitiva entre dos interacciones similares pueden constituir una forma común de regular la expresión de genes de proteínas que forman parte de aglomerados celulares o del complejo entre ácidos nucleicos y proteínas. Ya se conoce al menos otro caso similar.

Puesto que la represión de la traducción de proteínas ribosómicas depende de la disponibilidad de ARNr, importa saber cómo se controla la cantidad de ARNr que debe haber. Recordemos un aspecto interesante de la síntesis de ribosomas: su tasa de fabricación, que es la misma, aun cuando se utilicen distintos medios de cultivo. Por ejemplo, un medio pobre en fuente de carbono y rico en fuente de nitrógeno produce la misma cantidad de ribosomas en *E. coli* que un medio rico en carbono y pobre en nitrógeno, siempre que ambos medios ofrezcan condiciones igualmente favorables para el crecimiento. Lo mismo ocurre con muchas otras sustancias requeridas para el crecimiento celular.

El modelo del operón requiere la presencia de una señal química (inductor) para activar y reprimir los genes. El inductor es un derivado químico de una sustancia presente en el medio ambiente bacteriano. Hemos visto que un medio con una composición química

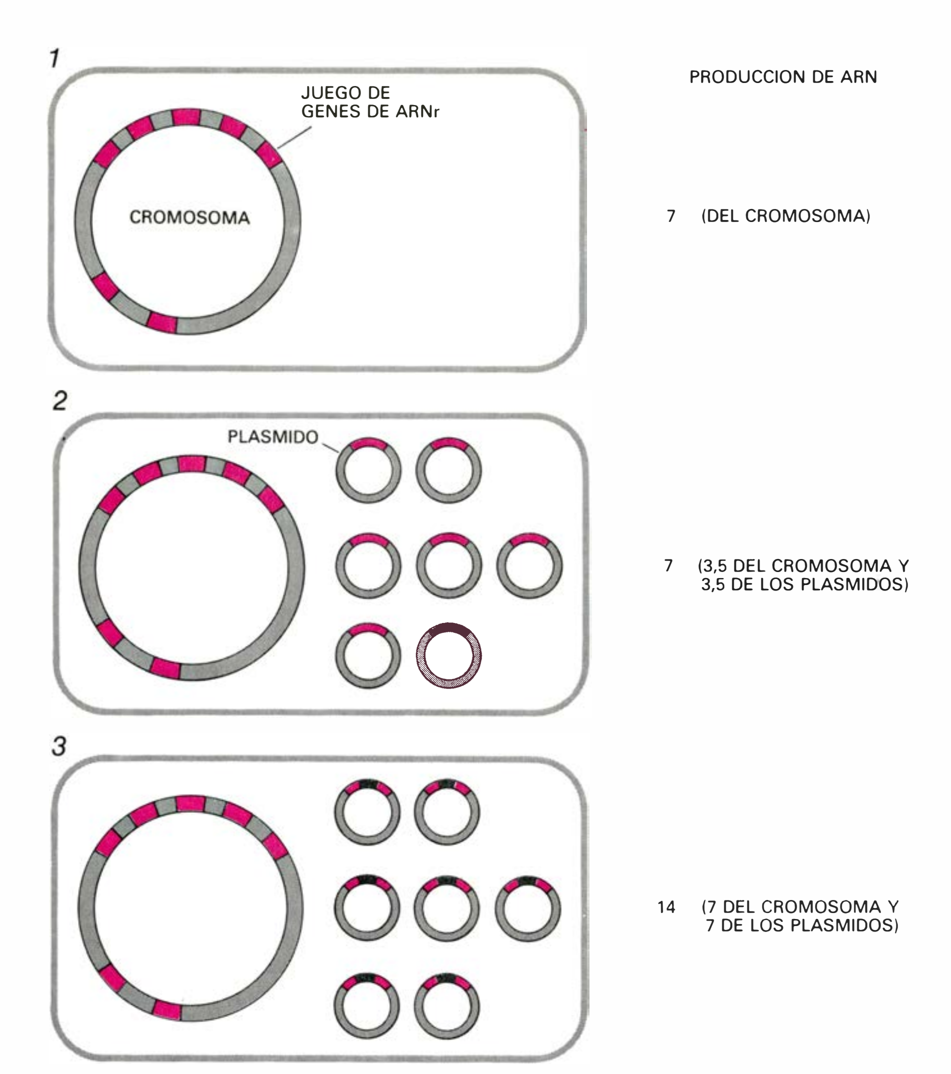
9. SITIOS DE UNION de los represores de la traducción, en el ARN ribosómico y el ARN mensajero. El esquema muestra semejanzas significativas. Durante el ensamblaje del ribosoma, la proteína *S8* forma un complejo con el ARNr 16*S*. También puede formar complejo con su propio ARNm. Una vez unida al ARNm, impide la traducción de varias proteínas de su operón, comenzando por *L5*. En ambos casos, la molécula de *S8* interacciona con una región bicatenaria de la cadena de ARN, cuya estructura ha podido determinarse. La correspondiente al ARNr 16*S* se muestra a la izquierda y, la del ARNm, a la derecha. Las regiones bicatenarias se forman cuando se emparejan bases complementarias (*líneas continuas*) o, de una forma más débil, *G* con *U* (*líneas discontinuas*). Se aprecian grandes similitudes entre algunas regiones de las secuencias (*zonas coloreadas*). Tales similitudes permiten que el represor se una a ambos ARN.

particular puede producir la misma tasa de síntesis de ribosomas que otro con una composición bastante distinta. Ello quiere decir que, si la síntesis de ARN ribosómico estuviese regulada por un mecanismo de tipo operón, ambos medios deberían suministrar un mismo compuesto químico, y en cantidades similares, que sirviera como señal en el proceso regulador; lo que parece improbable, aunque no imposible.

Más probable parece que un segundo mecanismo de retroalimentación determine la tasa de transcripción de los ARN ribosómicos. Supongamos que, en general, la célula de *E. coli* posee capacidad para elaborar más ribosomas de los que sintetiza en un momento cualquiera. Rara vez se pone a prueba la capacidad de la bacteria, puesto que los ribosomas libres (los que no estén fabricando proteína) suelen limitar la tasa de transcripción del ARNr y, por tanto, el ensamblaje de nuevos ribosomas.

Para sopesar esta hipótesis, Jinks-Robertson, Richard L. Gourse y el autor diseñaron un experimento de dosis génica para ARN ribosómico. Se introdujeron en *E. coli* plásmidos portadores de un juego completo de genes de ARNr. Dentro de la bacteria pueden coexistir varias copias del plásmido. En nuestro experimento había siete copias del plásmido por cada cromosoma bacteriano. Por tanto, cada célula transformada (que había captado el plásmido) tenía un total de 14 operones de ARNr, sus propios siete y los aportados por los plásmidos.

La duplicación del número de operones de ARN ribosómicos no incrementó, o lo hizo muy ligeramente, la tasa total de síntesis de ARNr. Los operones aportados por los plásmidos eran activos: su tasa de transcripción de ARNr era aproximadamente la misma que presentaban los genes del cromosoma bacteriano. Los 14 promotores de los ARNr funcionaban todos a la mitad de su tasa normal, con lo que la síntesis total de ARNr era la misma que presentaban las bacterias sin plásmidos.



10. UN SEGUNDO MECANISMO RETRORREGULADOR controla la transcripción del ARN ribosómico en *E. coli*. Aquí se esquematiza el experimento que permitió demostrar su existencia. Se cultivaron tres grupos de células de *E. coli* en el mismo tipo de medio. Los tres sintetizaban la misma cantidad de ribosomas. Uno de los grupos (*1*) llevaba sólo los siete juegos de genes de ARNr (*color*) presentes en el cromosoma bacteriano. Los ARNr transcritos a partir de ellos se incorporaban en los ribosomas. Los ribosomas libres, los que no estaban sintetizando proteínas, actuaban como represores, impidiendo la transcripción de más ARNr. En un segundo grupo de células se insertaron plásmidos (*2*) que llevaban más juegos de genes de ARNr (*color*). El segundo grupo de bacterias fabricaba la misma cantidad de ARNr que el primero, mitad procedente de los plásmidos y mitad del cromosoma. El ARNr de ambas fuentes se incorporaba en los ribosomas, que a su vez actuaban como represores. El tercer grupo de células también llevaba plásmidos (*3*), pero con deleciones en los genes de ARNr (*negro*). El ARNr procedente de los plásmidos no se incorporaba en los ribosomas; el ARN plasmídico no intervenía, pues, en la regulación de la producción de ARNr. Se comprobó que el tercer grupo fabricaba el doble de ARNr que los demás.

El resultado del experimento de dosis génica es coherente con el modelo retroactivo de la limitación de la síntesis de ARNr. De acuerdo con este modelo, la tasa de síntesis total de ARNr no excederá significativamente la tasa requerida para sintetizar el número óptimo de ribosomas. Si se sintetiza un exceso de ARNr, se acumulará en forma de ribosomas libres, que inhibirán la síntesis de nuevos ARNr, tanto por parte de los operones presentes en los plásmidos como de los del cromosoma.

Estos resultados, sin embargo, admiten otra explicación, si se presume la existencia de una molécula necesaria para la transcripción de ARN ribosómico. La cantidad del hipotético factor vendrá determinada por la composición del medio ambiente de la bacteria; no estará pues afectada por el número de juegos de genes de ARNr. A éste se le podría denominar mecanismo del factor limitante.

Para cribar ambas hipótesis, insertamos, en un plásmido, genes de ARNr a los que les habíamos eliminado previamente ciertos nucleótidos. Introdujimos los plásmidos en *E. coli*. Por efecto de las deleciones, los ARNr transcritos a partir de los operones presentes en los plásmidos perdieron su funcionalidad: no se integraban en los ribosomas. Por tanto, las bacterias portadoras de plásmidos tenían más genes de ARNr y promotores, pero sólo el número normal de genes capaz de sintetizar ARN que pudiera formar parte de un ribosoma.

Si la teoría del factor limitante era correcta, habría una cantidad total de ARN ribosómico sintetizada en las células transformadas igual a la cantidad sintetizada en las células con el comple-

mento normal de operones de ARNr: por la sencilla razón de que ambos grupos de células crecerían en el mismo medio. Con la misma tasa de crecimiento, los dos tipos celulares conseguirían idéntica cantidad del hipotético factor que limita la transcripción del ARNr.

Si hubiese un mecanismo de retrorregulación, el resultado sería diferente. El ARN ribosómico defectuoso, procedente de los operones de los plásmidos, no formaría parte de los ribosomas ni serviría para elaborar proteínas o inhibir la síntesis de más ARNr. Sólo el ARN funcional incorporado en ribosomas, que se acumularían luego como ribosomas libres, podría actuar como represor. Por tanto, las células con genes de ARNr defectuosos sintetizarían más ARNr, incluyendo tanto el ARNr defectuoso como el funcional, que las bacterias no alteradas.

El experimento de dosis génica demostró un notable incremento de la síntesis de ARN ribosómico en las células portadoras de los genes defectuosos, apoyando así la hipótesis de la retroinhibición. No obstante, se desconoce si el control de la síntesis de ARNr lo ejercen directamente los ribosomas libres o una molécula intermediaria formada en respuesta al incremento de ribosomas libres.

Aunque quedan por resolver algunos detalles, en líneas generales ya sabemos cómo se controla en *E. coli* la síntesis de ribosomas. En un medio favorable para el crecimiento de *E. coli*, la concentración de ribosomas es alta, y casi todos ellos sintetizan proteínas. Son pocos los ribosomas que quedan libres en la célula.

Si se transfieren las bacterias a un medio menos favorable para su crecimiento, disminuye la tasa de síntesis de proteínas, subiendo el número de ribosomas libres. Estos inhiben, directa o indirectamente, la síntesis de ARN ribosómico, así como la síntesis de otros ARN, incluidos los ARN de transferencia, que no desempeñan un papel importante en nuestra historia. La ausencia de ARNr recién transcrito determina que se interrumpa la síntesis de proteínas ribosómicas, por retroinhibición de la traducción. Finalmente, la concentración de ribosomas acaba gradualmente ajustándose a una menor tasa de crecimiento.

Si las células de *E. coli* que crecen en un medio pobre se transfieren ahora a otro que favorece el crecimiento, la tasa de síntesis de proteínas se acelera rápidamente. El incremento inmediato es pequeño, porque normalmente sólo una cantidad exigua de ribosomas está disponible para el trabajo de síntesis. Sin embargo, ese pequeño incremento en la síntesis de proteínas supone ya una gran reducción en la concentración de ribosomas libres, lo cual dispara la síntesis de ARN ribosómico y, por tanto, acelera la acumulación de ribosomas.

Todo ello concuerda con lo observado al transferir las bacterias a un medio más rico. Tras el cambio, la tasa de síntesis de ribosomas se incrementa más rápidamente que la de proteínas. La tasa de síntesis de proteínas crece inicialmente por la utilización de los pocos ribosomas libres; para que el incremento sea mayor hay que esperar a que se sinteticen nuevos ribosomas. Normalmente se alcanza un estado estable en el que la síntesis de ribosomas está en equilibrio con el volumen de síntesis de proteínas determinado por las nuevas y favorables condiciones para el crecimiento y división celulares.

Los dos sistemas de retroinhibición parecen explicar cómo *E. coli* ajusta la síntesis de ribosomas para lograr una determinada tasa de crecimiento. La presencia de una cantidad considerable de ribosomas libres en la célula, cuando el crecimiento es muy bajo, sugiere que, en general, el crecimiento celular no está limitado por la capacidad para sintetizar ribosomas. Por el contrario, parece que la bacteria ajusta la tasa de síntesis de ribosomas para lograr la mayor tasa de crecimiento posible en cada ambiente. El mecanismo utilizado por *E. coli* para controlar la tasa de síntesis de ribosomas debe haber conferido una considerable ventaja adaptativa a los antepasados de las líneas actuales de *E. coli*, cuando tales mecanismos evolucionaban en medios naturales competitivos.

Descifradas las líneas generales de la síntesis de ribosomas en una bacteria, el siguiente paso lógico será descubrir la regulación de dicha síntesis en células eucariotas. Por estar divididas en núcleo y citoplasma, el control de la síntesis de ribosomas en células eucariotas será algo más complicado que en bacterias. A pesar de tal complejidad, las células eucariotas, como las bacterias, parecen ajustar la tasa de síntesis de ribosomas de acuerdo con las condiciones ambientales para conseguir una tasa específica de crecimiento celular. Aunque se desconoce el mecanismo de regulación, una vez desentrañado el misterio de cómo *E. coli* lo consigue es posible que en un futuro no muy lejano conozcamos cómo controla la síntesis de ribosomas la célula eucariota.

Empaquetamiento de esferas

¿Cuál es, en el espacio, la más densa colocación de esferas? Mucho se ha progresado hacia la solución, sobre todo en 24 dimensiones. Los resultados tienen aplicación en la codificación numérica de información

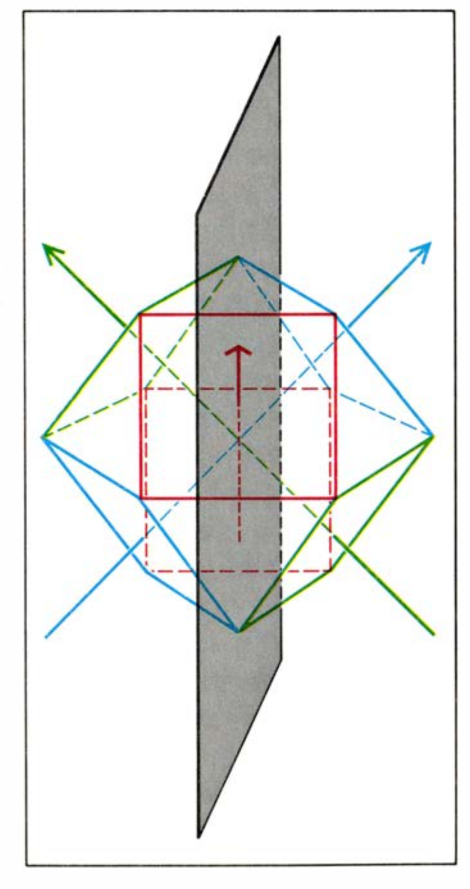
N. J. A. Sloane

A un fabricante de rodamientos de bolas le hacen un pedido por tantas bolas, todas del mismo tamaño, cuantas pueda expedir en el día a un puerto extranjero. Aunque las existencias de bolas son suficientes, tan sólo hay un barco disponible para el transporte. Su armador le explica al industrial que el calado del barco sería excesivo para la profundidad del canal del puerto si las bolas ocupasen más de las tres cuartas partes del volumen de las bodegas. El fabricante no se altera. "Su barco no correrá ningún peligro", le asegura. "Llene las bodegas hasta la escotilla". ¿Habrá de hacerle caso el armador al fabricante?

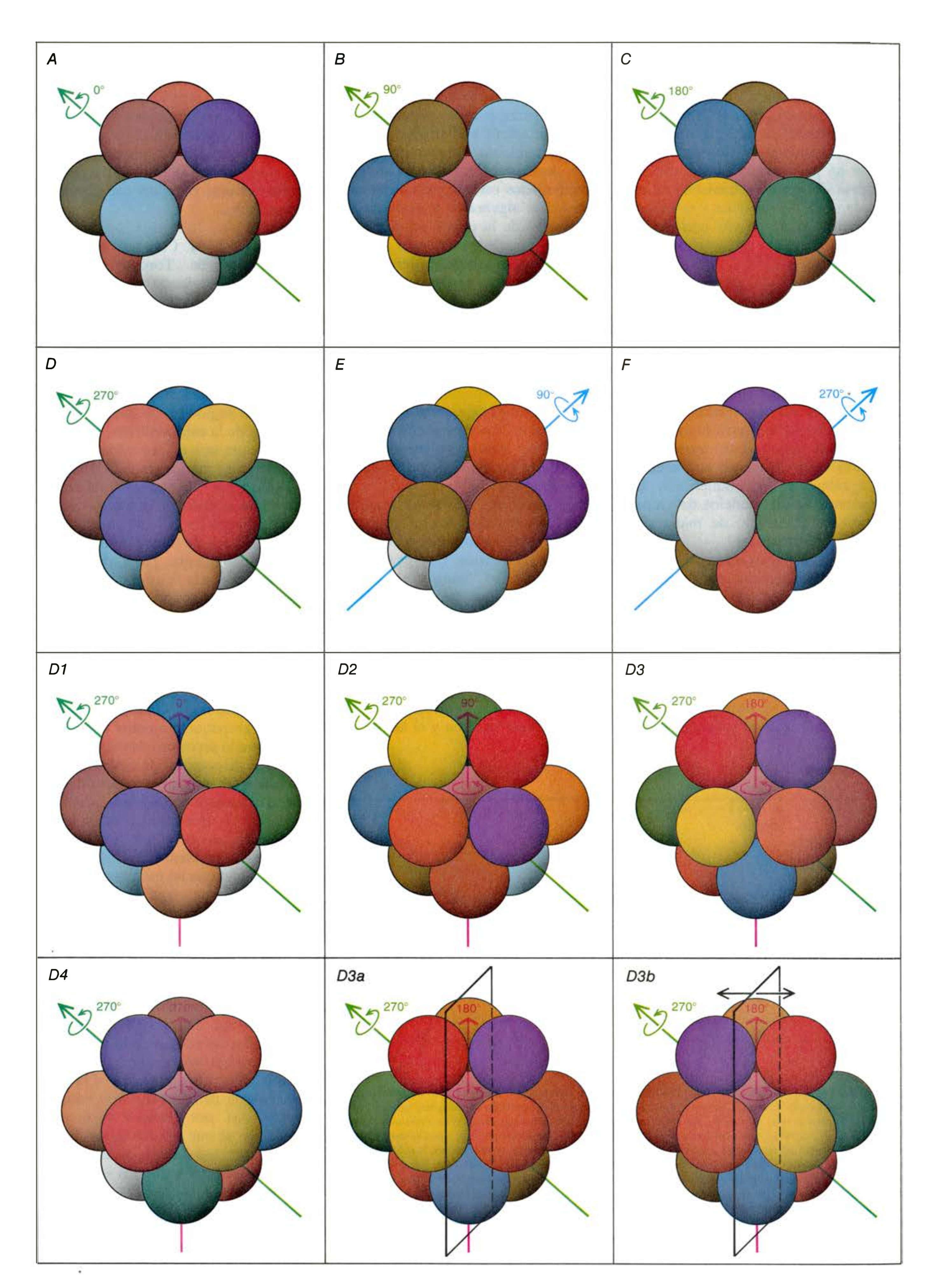
Para resolver el problema hay que averiguar cuán densamente se pueden empaquetar esferas en el espacio. Si en lugar de bolas para rodamientos, las bodegas del barco hubiesen de llenarse con cubos de igual tamaño, la respuesta sería fácil. Como los cubos se adosan perfectamente, sin dejar huecos ni intersticios (haciendo caso omiso de los pequeños espacios que pudieran quedar en torno a las paredes y el techo) las bodegas quedarían prácticamente repletas de cubos, y las seguridades dadas por el fabricante serían infundadas. Las bolas, sin embargo, no pueden ser empaquetadas sin pérdidas de espacio. Si a pesar de empacar las bolas tan densamente como sea posible el espacio desperdiciado excede todavía de la cuarta parte del volumen de las bodegas, el carguero puede llenarlas del todo y atravesar el canal sin peligro.

Bastan por lo general unos cuantos minutos de experimentación con naranjas o bolas de billar para convencer –erróneamente– a muchos de que el problema es trivial. Juntemos tres esferas sobre una superficie plana: sus centros formarán un triángulo equilátero; prosigamos añadiendo esferas sobre la superficie, de modo que cada nueva toque al menos a dos de las ya colocadas. Formemos una segunda capa de esferas alojando las nuevas sobre el "agujero" o depresión profunda que ha quedado en el centro de cada grupo triangular de esferas de la primera capa. Al acabar, la segunda capa es idéntica a la primera, si bien con un corrimiento en el plano horizontal. Añadiendo nuevas capas por el mismo procedimiento, el empaquetamiento de esferas resultante es el llamado empaquetamiento cúbico centrado en caras, que tan familiar es en química y cristalografía, y que ocupa apenas más del 74 por ciento del volumen de espacio. Hasta la fecha, éste es el más denso de los empaquetamientos conocidos.

Lamento tener que añadir que jamás ha podido demostrarse matemáticamente que este empaquetamiento sea maximal. La mínima cota superior para la densidad de empaquetamiento que se ha logrado hasta la fecha fue calculada por C. A. Rogers, de la Universidad de Birmingham, en 1958, quien demostró que un empaquetamiento de esferas nunca podrá tener densidad mayor de 0,7796 aproximadamente. Este resultado carece de utilidad particular para quienes busquen formas más eficientes de empaquetar bolas para rodamientos, pues la demostración de Rogers no ofrece ninguna construcción de modelos de empaquetamiento que se acerquen a su cota. (En el artículo donde daba a conocer su demostración, el



1. EMPAQUETAMIENTO CUBICO centrado en las caras; es tenido por el más denso de los empaquetamientos de esferas en el espacio tridimensional. Solemos verlo en apilamientos de frutas en los puestos del mercado y en las pilas de balas de cañón de ciertos monumentos. A pesar de siglos de esfuerzos, aún no es posible dar una demostración de su carácter maximalmente denso. Cada esfera del empaquetamiento "besa", o sea, es tangente, a otras 12 esferas; la demostración de que tal número es máximo no se consiguió hasta 1874. Dejando fijo el centro de una esfera, el conjunto de todos los giros y simetrías que permutan entre sí las 12 esferas circundantes es el llamado grupo de simetrías del empaquetamiento. Para el centrado en las caras, el grupo consta de 48 elementos; para comprenderlos mejor, resulta conveniente imaginar que el centro de cada esfera está situado en uno de los vértices del poliedro mostrado a la derecha, llamado cubo-octaedro. Cualquiera de las seis caras cuadradas del cubo-octaedro puede convertirse en la cara frontal haciendo girar adecuadamente el sólido en torno a los ejes azul o verde ($A-F$). Cada conjunto de cuatro esferas que formen una cara cuadrada (la cuarta cara, pongamos por caso) puede entonces adoptar una de cuatro configuraciones, haciendo girar la figura completa en torno al eje rojo (*D*1-*D*4). Finalmente, cada configuración (la tercera, por ejemplo) puede quedar reflejada respecto a un plano vertical, produciendo una nueva configuración (*D3a, D3b*). El número total de elementos del grupo de simetrías es, entonces, 6×4×2, es decir, 48. Un grupo de simetrías análogo, que describe los giros y simetrías de un empaquetamiento denso de esferas en el espacio de 24 dimensiones, descubierto por John Leech, de la Universidad de Glasgow, ha sido de gran importancia en la teoría matemática de grupos finitos.


A
0°
B
90°
C
180°
D
270°
E
90°
F
270°
D1
270°
0°
D2
270°
90°
D3
270°
180°
D4
270°
D3a
270°
180°
D3b
270°
180°

propio Rogers hacía notar que "muchos matemáticos están convencidos, y todos los físicos saben", que la solución correcta ronda en torno al 74 por ciento.) En el cuarto de siglo transcurrido desde el comentario de Rogers, su estimación ha permanecido invariable; el problema del empaquetamiento de esferas, tan fácil de enunciar y tan difícil de resolver, sigue siendo uno de los enigmas elementales no resueltos de las matemáticas.

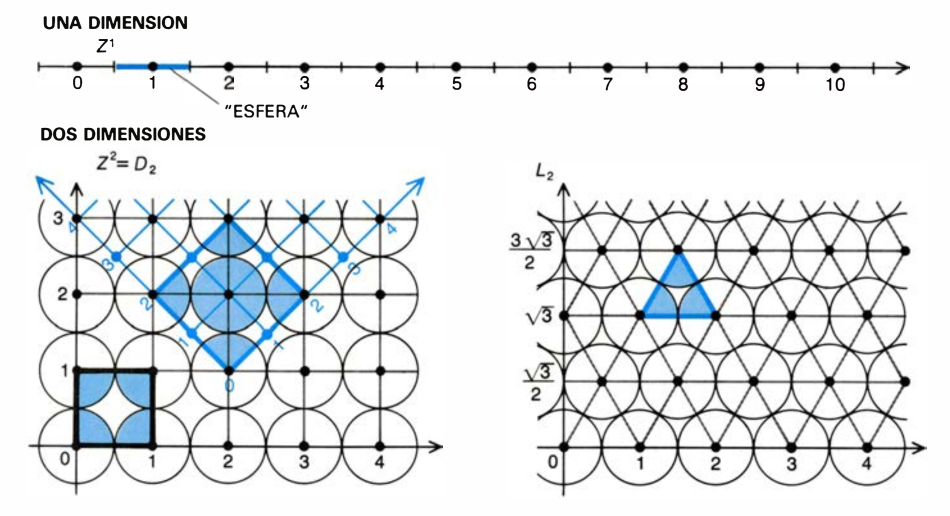
Durante muchos años han venido estudiándose los empaquetamientos densos de esferas; en parte, a causa de sus extensas connotaciones con el comportamiento de sólidos y líquidos. Las propiedades moleculares de muchos materiales cristalinos pueden describirse, al menos en primera aproximación, como resultantes de las diversas fuerzas que actúan sobre una ingente agrupación de esferas estrechamente empaquetadas. Igualmente importantes son las aplicaciones de los modelos de empaquetamiento al estudio de materiales porosos o pulverulentos.

A pesar del interés que revisten los estudios experimentales de los empaquetamientos de esferas a la hora de explicar ciertos sistemas físicos, existen igualmente razones de peso para llevar a cabo el estudio matemático de empaquetamientos esféricos idealmente densos. El mismo hecho de que no haya podido demostrarse para ningún empaquetamiento que es maximalmente denso hace pensar que la comprensión matemática del espacio euclídeo tridimensional ordinario dista de ser completa. Además, para los matemáticos la noción de esfera y el problema de su empaquetamiento pueden generalizarse, dando cabida a objetos matemáticos llamados esferas *n*-dimensionales cuya forma algebraica se asemeja a la descripción de la esfera en el espacio ordinario. Desde hace tiempo se ha reconocido que el estudio de empaquetamiento de esferas en espacios de *n* dimensiones es matemáticamente equivalente al diseño de un conjunto finito de mensajes codificados digitalmente, que no desperdicien energía ni causen confusión al ser transmitidos. En estos últimos años, la búsqueda de empaquetamientos óptimos de esferas en espacios de dimensión 24 o superior ha llevado a importantes descubrimientos en la rama de las matemáticas conocida por teoría de grupos.

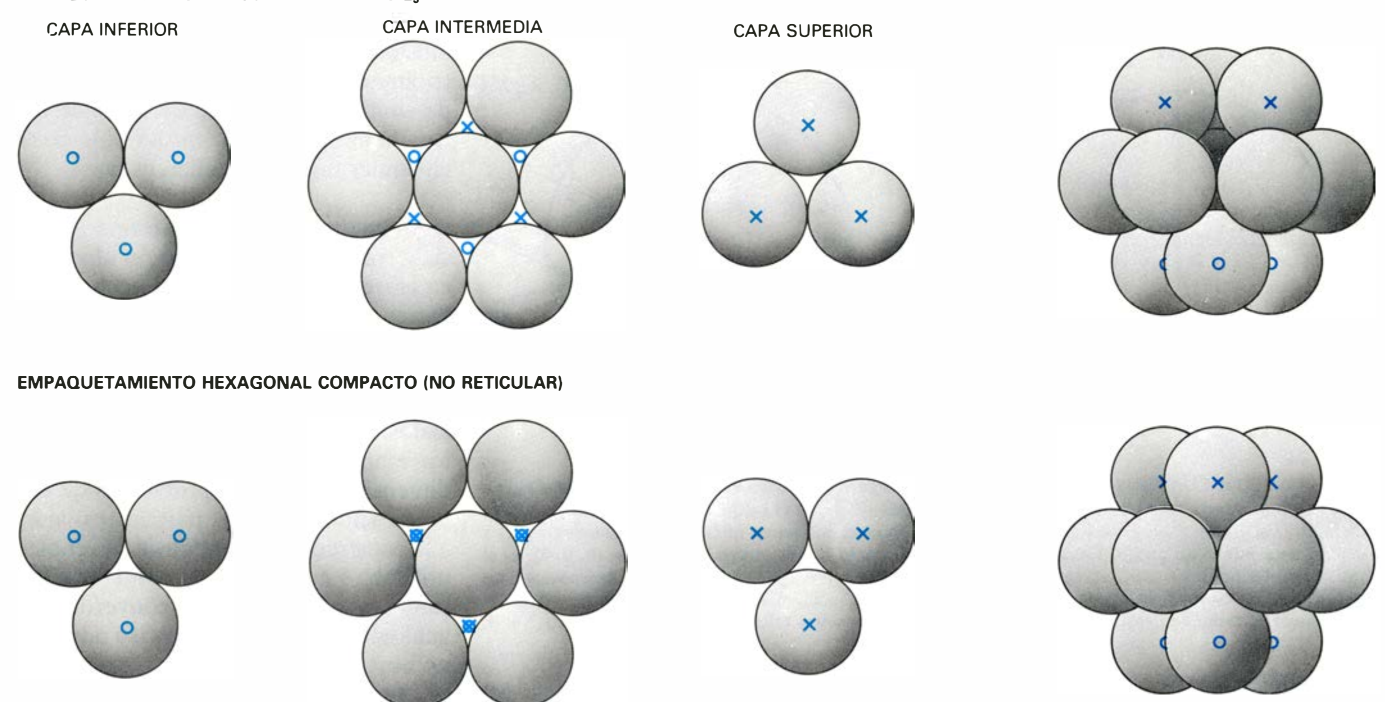
Hay otros dos problemas, íntimamente relacionados con el problema del empaquetamiento, que revisten importancia geométrica. Uno es el llamado problema del "número de osculación", a saber, ¿cuántas esferas pueden situarse en torno a una central de modo que todas las esferas circundantes "besen", es decir, sean tangentes en un punto, a la central? Para tres dimensiones, el problema de osculación fue motivo, en 1694, de una célebre controversia entre Isaac Newton y el astrónomo escocés David Gregory. Newton sostenía que el número de osculación era 12; para el empaquetamiento cúbico centrado en caras, antes descrito, tal número es efectivamente 12 [*véase la figura 1*]. Gregory aducía, probablemente, que podía forzarse la entrada de una décimotercera, pero no fue capaz de demostrarlo.

Según H. S. M. Coxeter, de la Universidad de Toronto, Gregory pudo haber imaginado que las doce esferas exteriores podían rodar en torno a la central de forma tal que todos los huecos pudieran concentrarse en una misma dirección, dejando de esta forma sitio para 13 esferas circundantes. Y en realidad, no es difícil demostrar que el ángulo sólido, medido desde el centro de la esfera núcleo, subtendido por cada una de las esferas circundantes es algo menor que 1/13 del ángulo sólido total. El volumen total del espacio que contornea la esfera central es, en principio, suficiente para acomodar el volumen de trece esferas. Empero, inevitablemente, parte del ángulo sólido visto desde la esfera central esta subtendido por los intersticios que se producen al empaquetar las esferas circundantes. La cuestión del número de osculación no quedó dirimida hasta 1874. R. Hoppe demostró que Newton tenía razón.

El segundo importante problema relativo a empaquetamientos de esferas es el llamado problema de recubrimiento. ¿Qué disposición de esferas idénticas consigue, con mínima densidad de empaquetamiento, que todo punto se encuentre en el interior o sobre la superficie de una de ellas? A diferencia del primer problema de empaquetamiento, en el que las esferas no pueden penetrar unas en otras, la solución del problema de recubrimiento exige que las esferas se solapen. Una forma de cubrir totalmente un volumen con esferas es inflar todas las de un empaquetamiento hasta que engloben en sí todos los orificios y pasadizos del empaquetamiento inicial. En general, sin embargo, al inflar las esferas de la más densa de las configuraciones de empaquetamiento conocidas no se obtiene la solución óptima del problema de recubrimiento. Así, por ejemplo, se tiene la convicción de que, en tres dimensiones, el recubrimiento óptimo es el dado por esferas dispuestas según los vértices de lo que se llama un retículo cúbico centrado en el cubo. Sin embargo, al tomar un empaquetamiento de esferas que no se intercepten mútuamente, con centros en los centros de los cubos, el em-



2. LOS EMPAQUETAMIENTOS DE "ESFERAS" pueden tener lugar no sólo en tres dimensiones, sino también en una y dos. En una dimensión, las esferas son segmentos rectilíneos de longitud unidad centrados en los puntos de abscisa entera. Tales esferas cubren el 100 por ciento de la recta; cada una "besa" a otras dos. Como es obvio, éste es el empaquetamiento de máxima densidad posible, y recibe la designación de Z^1. En el plano, las esferas son círculos, y existen tres empaquetamientos de interés. En el llamado Z^2, las esferas (los círculos) tienen sus centros en todos los puntos de coordenadas enteras. En el empaquetamiento D_2, las esferas tienen sus centros en puntos alternantes del empaquetamiento Z^2, por así decirlo, al tresbolillo. Dando a los ejes de coordenadas de Z^2 un giro de 45 grados, y redefiniendo la unidad de escala, resulta el empaquetamiento D_2. Ambos empaquetamientos son, por tanto, equivalentes. Su densidad es igual a la parte de cuadrado recubierta por círculos o sectores circulares (*sombreados en color*); es de $\pi/4$, aproximadamente 0,7854 (*izquierda*). El más denso de los empaquetamientos de círculos en dos dimensiones es reticular hexagonal, L_2. Los tres sectores circulares recubren $\pi\sqrt{3}/6$ del área total de cada triángulo equilátero. La densidad del empaquetamiento es, por tanto, de 0,9069, aproximadamente (*derecha*).



3. CONSTRUCCION del empaquetamiento reticular laminado L_3, equivalente al cúbico centrado en las caras D_3. Se efectúa adosando superpuestas capas de esferas cuyos centros se hallan situados de acuerdo con el retículo hexagonal L_2. Empaquetando las esferas de modo que las situadas en la tercera capa hexagonal se encuentren directamente encima de las correspondientes a la primera capa, el empaquetamiento resultante es el llamado hexagonal compacto, que tiene la misma densidad que L_3; los centros de las esferas del empaquetamiento no cumplen las condiciones de la definición de retículo.

paquetamiento resultante no es tan denso como otros conocidos, verbigracia, el centrado en las caras. Por otra parte, la conjetura de que los empaquetamientos centrados en el cubo resuelven el problema de recubrimiento tampoco ha sido demostrada.

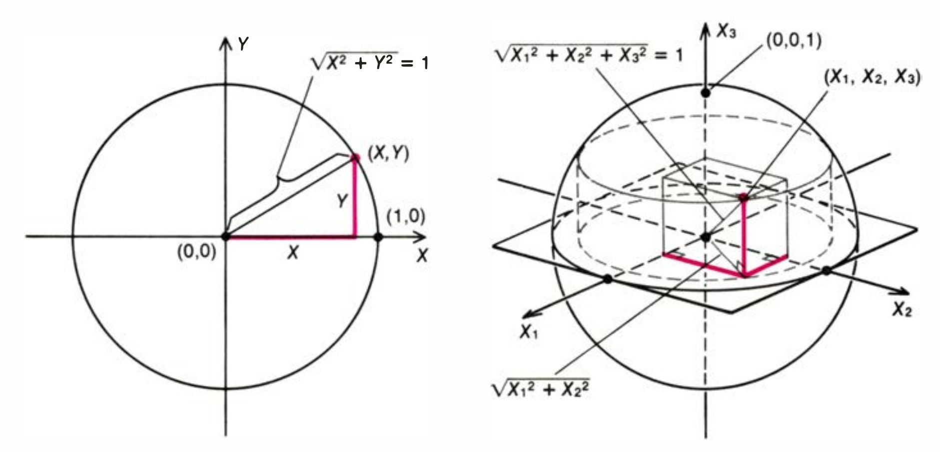
¿En qué reside la dificultad que impide resolver el problema de los empaquetamientos tridimensionales de esferas? Volvamos a fijarnos en el empaquetamiento centrado en las caras. Dado que el problema tan sólo estipula que ha de hacerse máxima la densidad del empaquetamiento, el número de esferas a empaquetar debe considerarse ilimitado. Retirando, pongamos por caso, un millón de esferas del empaquetamiento, la densidad permanecerá sin variación efectiva. Dicho de otra forma, se puede hacer que el cambio de densidad sea tan pequeño como se quiera aumentando el número de esferas del empaquetamiento considerado como un todo. Por otra parte, cada vez que se retira del empaquetamiento cúbico centrado en caras una esfera, se crea un empaquetamiento distinto. Así pues, toda demostración de que el empaquetamiento centrado en las caras es optimal (como tanta gente cree) es al mismo tiempo una demostración de que hay infinito número de distintos empaquetamientos igualmente optimales.

Hay otra forma de percibir las complejidades del problema de empaquetamiento de esferas en el espacio ordinario. En tres dimensiones, a lo sumo cuatro esferas pueden ser mutuamente tangentes. (Es preciso distinguir bien esta propiedad de la configuración exigida por el problema del número de osculación, donde las esferas circundantes no tienen por qué ser tangentes entre sí.) Cuando cuatro esferas se encuentran en mutuo contacto, sus centros ocupan los vértices de un tetraedro regular, una pirámide cuyas caras son triángulos equiláteros. Dado que es imposible aproximar más en el espacio las cuatro esferas, la configuración más densa que pueden adoptar cuatro esferas es la tetraédrica.

Supongamos que se va añadiendo de una en una nuevas esferas a una configuración tetraédrica, buscando formar otra configuración tetraédrica, siempre que sea posible. De poderse continuar indefinidamente este procedimiento, la disposición de esferas resultante tendría la máxima densidad posible, pues cada configuración tetraédrica se encontraría tan densamente empaquetada cuanto es posible. La cota superior dada por Rogers se basa precisamente en este razonamiento; un cálculo directo de trigonometría esférica muestra que alrededor del 0,7796 del volumen del tetraedro está ocupado por las cuatro esferas tangentes con centros en sus vértices. (El valor exacto es $\sqrt{2}$ [3 arc cos(1/3) − π].) Por desgracia, los tetraedros no se adosan perfectamente para llenar un espacio. La estrategia de ir colocando esferas en formaciones tetraédricas siempre que así es posible –el que podríamos llamar "algoritmo ansioso"– acaba obligándonos a dar un paso en falso. En una cierta etapa del procedimiento la configuración de esferas presenta una superficie tal que resulta imposible añadir nuevas esferas sin desperdiciar espacio interior. Por tanto, aunque el algoritmo ansioso engendra empaquetamientos óptimos a pequeña escala, de unos cuantos diámetros esféricos, resulta que a escala global este algoritmo origina un empaquetamiento menos denso que el cúbico centrado en las caras.

Para poder avanzar en los tres problemas de empaquetamiento explicados, los matemáticos han encontrado útil reforzar la intuición geométrica con una representación analítica de las esferas por medio de coordenadas rectangulares cartesianas. Como es bien sabido, para especificar un punto cualquiera del plano bastan dos coordenadas, una horizontal, o abscisa, x, y otra vertical, u ordenada, y. El punto se denota entonces mediante un par ordenado (x, y). Por ejemplo, el par (3, 4) denota el





4. UN PUNTO QUEDA DETERMINADO en dos dimensiones asignando valores a dos coordenadas, x e y. Una circunferencia de radio 1, con centro en el origen (0, 0), es el conjunto de los puntos (x, y) que satisfacen la ecuación $x^2 + y^2 = 1$ (*izquierda*). En tres dimensiones hacen falta tres coordenadas, x_1, x_2 y x_3, para determinar un punto. La superficie esférica de radio 1, con centro en el origen (0, 0, 0), es el conjunto de los puntos (x_1, x_2, x_3) que satisfacen la ecuación $x_1^2 + x_2^2 + x_3^2 = 1$. En n dimensiones, los puntos quedan especificados por n coordenadas $x_1, x_2, \ldots, x_n$. La "superficie" de una esfera n-dimensional de radio 1 y centro en el origen es el conjunto de puntos tales que $x_1^2 + x_2^2 + \ldots + x_n^2 = 1$.

punto del plano situado a tres unidades a la derecha del origen, en la dirección del eje x, y cuatro unidades por encima del origen, según el eje y.

La distancia entre el punto (3, 4) del plano y otro punto (x, y) cualquiera puede calcularse gracias al teorema de Pitágoras para triángulos rectángulos: el cuadrado de la distancia entre los dos puntos es igual al cuadrado de la distancia que los separa según el eje x, que es $(x - 3)^2$, más el cuadrado de la distancia medida a lo largo del eje y, $(y - 4)^2$. Dado que una circunferencia es, por definición, el conjunto de todos los puntos del plano equidistantes de un punto (a, b) central, todo punto (x, y) situado en ella ha de verificar la relación $(x - a)^2 + (y - b)^2 = R^2$, siendo R el radio de la circunferencia. En el caso de que el radio sea igual a 1, y el centro se encuentre en el origen (0, 0), la ecuación es mucho más sencilla: los puntos de la circunferencia han de satisfacer la ecuación $x^2 + y^2 = 1$.

Análogamente, los puntos del espacio tridimensional quedan determinados por tres coordenadas x, y, z, o con notación más sugestiva (x_1, x_2, x_3). La superficie esférica de radio 1, con centro en el origen de coordenadas, está formada, pues, por el conjunto de todos los puntos (x_1, x_2, x_3) tales que $x_1^2 + x_2^2 + x_3^2 = 1$, ecuación que emana, lo mismo que en dos dimensiones, de la definición geométrica de la esfera y de aplicar por dos veces el teorema de Pitágoras.

Cuando el número de dimensiones es mayor que tres, de poco sirve la intuición geométrica, y se hace imprescindible comenzar a pensar a partir de coordenadas. Un "punto", por ejemplo del espacio cuatridimensional, es un objeto matemático que exige cuatro números reales para quedar determinado sin ambigüedad; un tal punto se denota (x_1, x_2, x_3, x_4). Si la estatura, peso, edad e ingresos de una persona fuesen suficientes para hallar inequívocamente su nombre en una lista, estas cuatro cantidades podrían ser consideradas como determinantes de un punto en un espacio tetradimensional.

Las esferas de cuatro dimensiones se definen por analogía con las definiciones de circunferencia y esfera en dos y tres dimensiones. Todos los puntos (x_1, x_2, x_3, x_4) de la "superficie" de la esfera se hallan a una cierta distancia R de un punto central (a_1, a_2, a_3, a_4). La suma de los cuadrados de las distancias que separan según cada uno de los ejes de coordenadas a un punto (x_1, x_2, x_3, x_4) situado en la esfera del centro (a_1, a_2, a_3, a_4) ha de ser siempre igual a R^2.

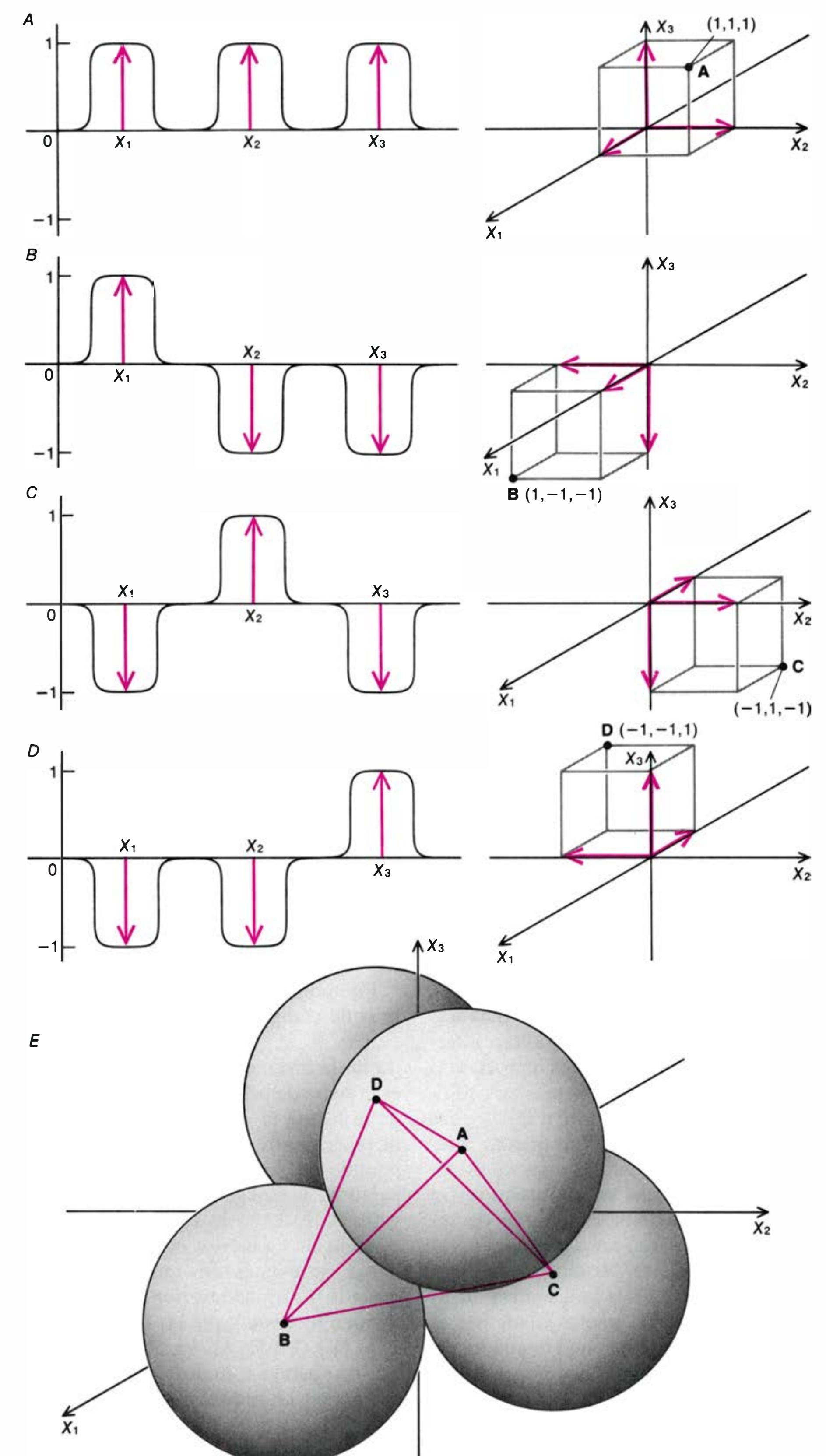
Se han escrito grandes tonterías acerca de los misterios de la cuarta dimensión. En los análisis matemáticos, no debe suponerse, como acostumbra a hacerse en física, que la cuarta dimensión representa el tiempo. Es preciso, además, evitar la tentación de reificar los términos, un tanto metafóricos, "superficie", "punto", "esfera", etcétera, que son aplicados a objetos matemáticos de propiedades algebraicas perfectamente válidas. La justificación del uso de tales términos está en que los objetos a que aluden han sido construidos por analogía con las propiedades algebraicas de esferas y circunferencias ordinarias; no debe suponerse, empero, que tales objetos representen objetos geométricos reales de algún universo más amplio que el nuestro. Repitámoslo: en matemáticas, el espacio cuatridimensional está formado por puntos de cuatro coordenadas en lugar de las tres habituales, y lo mismo vale para cualquier número de dimensiones.

En cuanto la noción de esfera queda traducida a una relación algebraica entre coordenadas, se descubre que varios problemas aparentemente desligados de los empaquetamientos de esferas pueden reducirse a uno de los tres problemas de empaquetamiento explicados. Las principales aplicaciones de los empaquetamientos en espacios de dimensión alta son a problemas de comunicación digital y, particularmente, a la construcción de sistemas de señales utilizables en canales de comunicación ruidosos y al diseño de convertidores analógico-digitales. Es preciso admitir, no obstante, que casi todos los análisis de estas aplicaciones han sido hasta ahora de carácter teórico; tan sólo muy recientemente se han utilizado empaquetamientos de esferas para diseñar sistemas prácticos. Conforme los sistemas de comunicaciones vayan haciéndose más y más refinados, se extenderá el ámbito de aplicaciones.

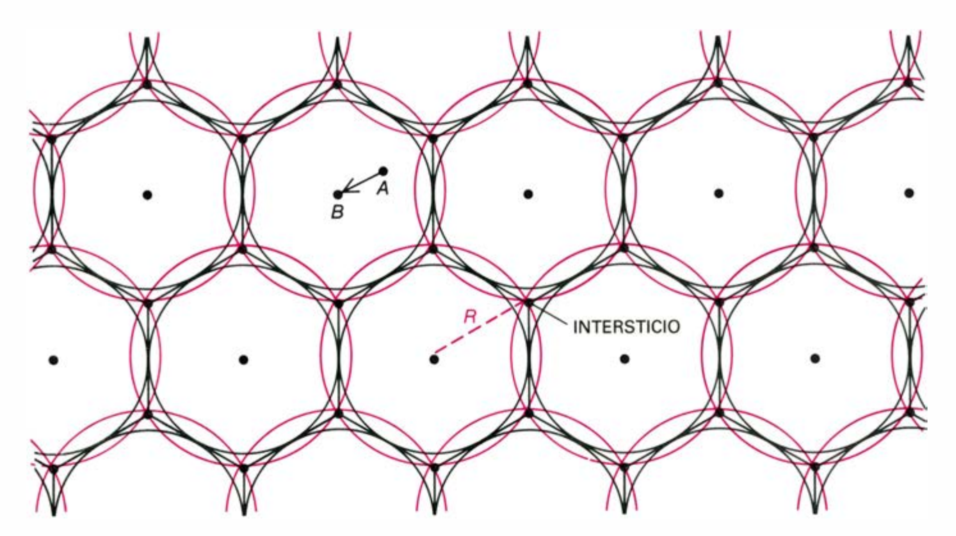
En las comunicaciones digitalmente codificadas, uno de los principales objetivos del diseñador es construir una lista de símbolos codificados distintos, llamados palabras de código, que poder transmitir con máxima fiabilidad y potencia mínima. Cada palabra-código podría venir representada, por ejemplo, por un símbolo de ocho caracteres, cada uno de los cuales puede tomar uno de cinco valores distintos: 0, ½, 1, −½, −1. A primera vista, parece que el sistema podría proporcionar un vocabulario de 5^8, o sea, 390.625 palabras de código distintas; pero las diferencias entre muchos pares de estas palabras de código son tan pequeñas que un tal sistema estaría muy expuesto a errores aleatorios en la transmisión, o a interferencias eléctricas. Por ejemplo, la diferencia entre la palabra de código (1, 1, 1, 1, 1, 1, 1, 1) y la (1, 1, 1, 1, 1, 1, 1, ½) es, con mucho, demasiado pequeña. Si se utilizasen ambas palabras de código, es posible que se produjeran frecuentes confusiones. Otra forma de hacer ver lo mismo es que si la diferencia entre dos palabras del código fuese tan pequeña como la diferencia entre (1, 1, 1, 1, 1, 1, 1, 1) y (1, 1, 1, 1, 1, 1, 1, ½), sería necesaria una gran potencia para garantizar que estas dos palabras se distinguiesen en presencia de ruido de fondo.

Existe una relación matemática de carácter general entre la discernibilidad de las palabras de código y la potencia necesaria para transmitirlas fidedignamente. Tal relación fue enunciada en 1948 por Claude E. Shannon, a la sazón en los laboratorios Bell, en su artículo *A Mathematical Theory of Communication*. Como afirma mi distinguido colega David Slepian, "quizá no se haya producido en este siglo otra obra que por sí sola haya alterado más profundamente la idea que el hombre tiene de la comunicación" que este artículo de Shannon. Lo que Shannon demostró es que dada una cantidad fija y finita de potencia, existe siempre un sistema de palabras de código que es posible transmitir esencialmente sin error. El único requisito es que la tasa o velocidad de transmisión de las palabras del código no rebase un umbral crítico llamado capacidad del canal de transmisión. Desdichadamente, el teorema de Shannon no es constructivo: demuestra la existencia de tales sistemas de codificación de señales, pero no da indicación de como podrían ser diseñados. Aunque se han construido muchos sistemas de codificación de señales, no se ha descubierto ninguno de rendimiento tan perfecto como el prometido por el teorema de Shannon.

Una forma de diseñar sistemas de señales que se aproximen a lo estipulado por el teorema de Shannon consiste en representar cada señal como un punto de un espacio n-dimensional. Fijémonos, por ejemplo, en una cualquiera de las sucesiones de ocho números del sistema de señales antes descrito. En términos físicos, cada uno de los números se corresponde con un determinado voltaje en la línea de transmisión; y por tanto, cada palabra del código puede representarse gráficamente en dos dimensiones, dibujando una serie de ocho impulsos cuya altura está especificada para cada uno de los ocho intervalos de tiempo establecidos en ese eje. Matemáticamente, empero, un único



5. DISEÑO DE CODIGOS para la eficiente transmisión de información; tiene que ver con los empaquetamientos de esferas. El código ha de ser un conjunto finito de señales, llamadas palabras de código, que sean fácilmente distinguibles entre sí y en cuya transmisión no se malgaste potencia eléctrica. Si cada palabra de código fuese, pongamos por caso, una secuencia de tres niveles discretos de tensión eléctrica, cada secuencia podría representarse por un punto del espacio tridimensional, cuya primera coordenada fuese el valor numérico del primer nivel de voltaje, la segunda, el valor del segundo nivel de tensión, y así sucesivamente (*A*−*D*). La potencia eléctrica requerida para la transmisión de cada impulso de tensión es proporcional al cuadrado de la tensión; por tanto, la potencia eléctrica total necesaria para transmitr una palabra de código es la suma de los cuadrados de las tres tensiones discretas asociadas a ella. Tal suma puede interpretarse como el cuadrado de la distancia que media entre el origen y el punto del espacio que representa geométricamente la palabra de código. Por tanto, el problema de hacer mínima la potencia de transmisión equivale al de situar todos los puntos representantes de las palabras del código tan cercanos al origen cuanto sea posible. Por otra parte, la necesidad de distinguir unas de otras a las palabras del código puede traducirse en el requisito de que los puntos que las representan en el espacio no se encuentren unos de otros a distancia menor que un mínimo *d* predeterminado. El cumplimiento simultáneo de ambas condiciones equivale a empaquetar esferas de radio *d*/2 tan densamente como sea posible (*E*).



6. CUANTIZACION DE DATOS dimanados de una fuente que varía en forma continua. Este problema está estrechamente ligado al de recubrir el espacio con una disposición de esferas mutuamente solapantes, con densidad mínima. Por ejemplo, para cuantizar datos en dos dimensiones, se procede primero a emparejar de dos en dos los datos de entrada; cada par de datos se trata cual coordenadas de un punto del plano. Cada punto de datos, como el *A*, por ejemplo, se redondea entonces a un punto de cuantización, tal como el *B*, yacente en una misma región predeterminada del plano que el punto de datos. El problema consiste en elegir los puntos de cuantización y la partición del plano de modo que el error medio de cuantización sea mínimo. Si los datos están uniformemente distribuidos, y como puntos de cuantización se toman los centros de una pavimentación por losetas cuadradas iguales, el error medio será 1/12. Puede deducirse un procedimiento de cuantización mejor gracias al recubrimiento óptimo del plano por medio de círculos. Tal recubrimiento es generado a partir del empaquetamiento hexagonal de circunferencias (*círculos negros*), incrementando el radio de cada círculo lo suficiente para que todo punto del plano quede contenido en el interior o la circunferencia de uno de los círculos (*círculos rojos*), por lo menos. El radio *R* de los círculos del recubrimiento es la distancia desde el centro hasta el "intersticio" más cercano del empaquetamiento. Conectando adecuadamente por segmentos rectilíneos los centros de las cavidades, el plano queda descompuesto en una partición de hexágonos regulares. Tomando como puntos de cuantización los centros de estos hexágonos, el error medio de cuantización, para datos uniformemente distribuidos, será de 5 $\sqrt{3}/108$, aproximadamente 0,0802, que viene a resultar ligeramente menor que 1/12.

punto de un espacio de ocho dimensiones permite representar la misma información: se toma como valor de la primera coordenada del punto el primer número de cada secuencia; para la segunda, el segundo, y así sucesivamente. Dado que un punto queda determinado en el espacio de ocho dimensiones al dar los valores de sus ocho coordenadas, cada una de las palabras del código del sistema puede ser representada por un único punto del espacio octodimensional.

Al representar por puntos las palabras de código podemos donar interpretación geométrica a dos importantes características de los sistemas de codificación. Recordemos, ante todo, que las palabras de código deben quedar fidedignamente distinguidas unas de otras. Lo cual sugiere que los puntos que en el espacio octodimensional hayan de representarlas deben encontrarse separados por una cierta "distancia" mínima. ¿Hasta qué punto es la distancia euclídea entre pares de puntos, dada por el teorema de Pitágoras, una buena medida de la discernibilidad de pares de palabras de código? Para determinar la distancia euclídea es preciso elevar al cuadrado las diferencias de valores de las coordenadas de los puntos. Al proceder así, las diferencias pequeñas entre valores de las coordenadas (es decir, las diferencias menores de una unidad) quedan reducidas, mientras las diferencias grandes (las mayores que 1) resultan amplificadas. Puesto que es mucho más probable que sean las pequeñas –y no las grandes– diferencias de voltaje las que produzcan confusiones en las palabras de código, la distancia calculada por el teorema de Pitágoras parece una medida razonable de la discernibilidad.

Fijémonos, por ejemplo, en las palabras de código (1, 1, 1, 1, 1, 1, 1, 1) y (½, ½, ½, ½, ½, ½, ½, ½). El cuadrado de la distancia entre los dos puntos es la suma de ocho cuadrados, cada uno de la forma $(1 - ½)^2$, y, por tanto, la distancia entre ellos es $\sqrt{2}$. Según la medida euclídea de distancias, las dos palabras de código (1, 1, 1, 1, 1, 1, 1, 1) y (0, 0, 1, 1, 1, 1, 1, 1), que sólo se diferencian en las dos primeras coordenadas, se encuentran también separadas por $\sqrt{2}$ unidades; por tanto, se distinguen lo mismo que las (1, 1, 1, 1, 1, 1, 1, 1) y (½, ½, ½, ½, ½, ½, ½, ½).

Un segundo e importante aspecto a considerar en todo sistema de palabras de código es que la potencia necesaria para transmitirlas ha de ser la menor posible. Se deduce de las definiciones elementales de potencia y voltaje eléctricos que la potencia es directamente proporcional al cuadrado del voltaje; en circuitos simples, la potencia es sencillamente el cuadrado del voltaje, dividido por la resistencia del circuito. La potencia total necesaria para trasmitir una palabra–código de ocho dígitos es, por consiguiente, la suma de los cuadrados de los ocho dígitos que la componen. Esta suma constituye el cuadrado de la distancia entre el punto del espacio octodimensional representante de la palabra de código y el origen (0, 0, 0, 0, 0, 0, 0, 0).

El diseño de un sistema de señales fidedigno, que utilice eficientemente la potencia disponible para la transmisión, puede así reducirse al problema geométrico de situar puntos en el interior de una región del espacio, con el requisito de que no se encuentren demasiado próximos. Si los puntos deben distar como mínimo $\sqrt{2}$ unos de otros, el problema es equivalente al de hallar el más denso empaquetamiento de esferas de radio igual a la mitad de esa distancia, o sea, $\sqrt{2}/2$. Un problema cercanamente emparentado es el de hallar un sistema de palabras de código que sean todas de la misma energía; problema que es equivalente al de situar tantos puntos cuantos sea posible sobre la superficie de una esfera *n*-dimensional, imponiéndoles al mismo tiempo la condición de que no se encuentren demasiado próximos. Este problema es, a su vez, una versión generalizada del problema del número de osculación de una esfera.

Resulta que en espacios octodimensionales sí existe un empaquetamiento extremadamente denso, llamado empaquetamiento E_8, que fue descubierto por los matemáticos rusos Alexander N. Korkin y E. I. Zolotareff y por un jurista aficionado a la matemática británico, Thorold Gosset, en el último tercio del siglo pasado. Los centros de las esferas del empaquetamiento E_8 son todos los puntos cuyas coordenas sean números enteros, o números enteros más un medio. La suma de las coordenadas de cada punto tiene que ser un número par. Existen 240 de tales puntos situados a distancia $\sqrt{2}$ del origen: 112 puntos son de la forma (±1, ±1, 0, 0, 0, 0, 0, 0), pudiendo los unos y cualquier combinación de signos presentarse en cualesquiera posiciones; más 128 puntos de la forma (±½, ±½, ±½, ±½, ±½, ±½, ±½, ±½), donde el número de signos menos es par.

El empaquetamiento E_8 pudiera convertirse en la base de un esquema práctico y eficiente de transmisión de seña-

les. Si el sistema hubiera de contener exactamente 240 palabras de código, podrían elegirse los 240 puntos del empaquetamiento E_8 que equidistan del origen. Empero, casi todos los sistemas prácticos exigen que el número de palabras de código a utilizar sea potencia entera de 2. Por ejemplo, para transmisiones telefónicas a distancias medianas se emplea mucho el sistema de transmisión digital conocido por modulación codificada por impulsos. La tensión eléctrica generada por la voz se mide cada 1/8000 de segundo; cada una de las medidas de muestra tomadas es cuantizada y sustituida por uno de $2^8 = 256$ niveles. Cada una de las tensiones cuantizadas es expresada entonces por un número binario de ocho dígitos, u octeto; el conjunto de tales números binarios componen el llamado código fuente.

Desde los trabajos de Harry Nyquist, de los laboratorios Bell, allá por los años veinte, se sabe que la señal de voz es reconstructible a partir, exclusivamente, de los valores de las muestras. La señal sonora puede ser imaginada como una magnitud continuamente variable con el tiempo, como lo son la presión del aire y la tensión eléctrica. A principios del siglo pasado, Jean Baptiste Joseph Fourier demostró que la gráfica de cualquier magnitud variable puede ser aproximada, con tanta precisión cuanta se desee, superponiendo curvas senoidales y cosenoidales de frecuencias y amplitudes apropiadas. Las curvas seno y cosenoidales (cuyo número puede ser infinito para curvas arbitrarias) se llaman componentes de Fourier de la curva de la gráfica.

Supongamos que la gráfica de una magnitud pudiera ser exactamente generada superponiendo un número finito de componentes de Fourier cuyas frecuencias no excedan de una cierta frecuencia W ciclos por segundo. Lo que Nyquist demostró es que entonces la gráfica podría también ser reconstruida exactamente a partir de los valores que toma cada $1/2W$ de segundo. Por ejemplo, una señal de voz que carece de componentes de frecuencia superior a 4000 ciclos por segundo puede ser perfectamente reconstruida a partir de muestras tomadas cada 1/8000 de segundo. Basta, por tanto, ir tomando muestras de la señal de voz, y transmitir solamente los valores de las muestras, que vienen representadas por las palabras del código fuente, y no transmitir la señal íntegra de voz. Si los valores de las muestras fuesen transmitidas sin redondeo previo, el teorema de Nyquist demuestra que todo un segundo de señal telefónica puede ser representado como un único punto de un espacio de dimensión 8000. He aquí la potencia de las matemáticas.

Para que la transmisión de los números del código fuente se haga con la máxima eficiencia, los valores tomados como muestra deben codificarse, ulteriormente, por el llamado código de canal. Es el código de canal el que tiene relación con el problema de empaquetamiento de esferas. Un excelente método para deducir un código de canal a partir del empaquetamiento E_8 consiste en codificar cada par de octetos binarios del código fuente. A cada uno de los números binarios de 16 dígitos resultantes se le asigna el punto central de una de las $2^{16} = 65.536$ esferas del empaquetamiento E_8; se construye así un buen código de canal escogiendo 65.536 puntos centrales que queden próximos al origen. En el extremo receptor de la línea telefónica, las palabras de código correspondientes a las coordenadas de cada punto central son reconvertidas en los números binarios del código fuente, y la señal sonora, reconstruida a partir de estos números binarios.

Existe una segunda importante aplicación de los empaquetamientos de esferas a las comunicaciones por codificación digital. La describiré sucintamente. Recordemos que al deducir de la señal telefónica los números binarios del código fuente era necesario cuantizar la intensidad verdadera de la señal, redondeándola y aproximándola a uno de 256 niveles. El mundo real está repleto de valores tan desgarbados como 0,7913..., pero el mundo de los ordenadores y de los sistemas digitales tiene, en última instancia, que trabajar solamente con valores "perfectos" como 0 y 1. Todo dispositivo destinado a redondear los valores de magnitudes continuamente variables y asignarles uno de los valores de un cierto conjunto discreto se llama convertidor de analógico en digital, o brevemente, cuantizador.

La cuantización puede efectuarse tanto en dos o más dimensiones como a lo largo de un único eje de coordenadas. Imaginemos el plano descompuesto en regiones, de formas no necesariamente congruentes, e imaginemos que en cada región haya sido señalado un punto. Cualquier formación de puntos y regiones de esas características es un cuantizador bidimensional; los datos que ingresan en el cuantizador son un par de números reales, que especifican un punto arbitrario; la salida es el punto de cuantización preseleccionado yacente en la misma región del plano que el punto arbitrario. Así pues, cualquier punto del plano queda "redondeado" a uno de los puntos de cuantización. El proceso, por así decirlo, comprime los datos de entrada; en lugar de los valores exactos de las coordenadas del punto-dato bastará transmitir un número-índice del punto cuantizador.

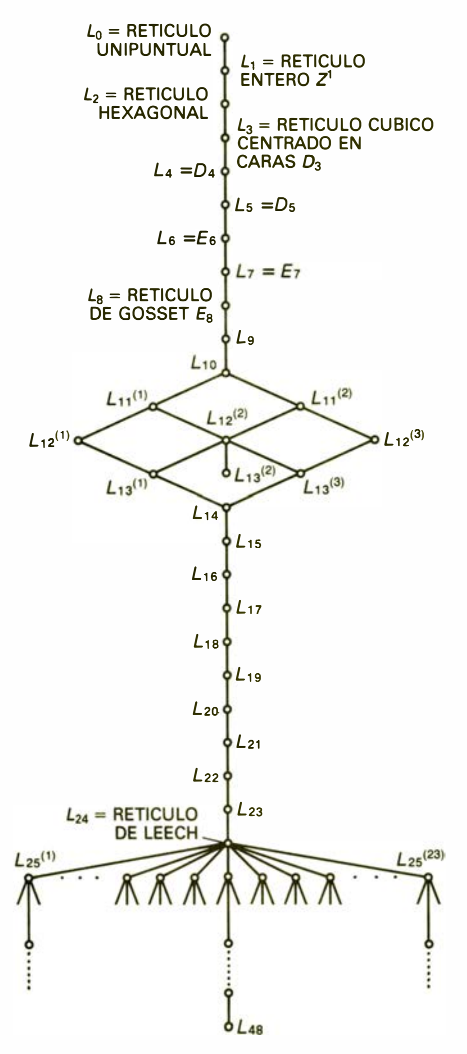
La cuantización introduce errores. Se busca, por tanto, elegir los puntos de cuantización de modo que el error promedio sea mínimo. Por ejemplo, si la distribución de valores de entrada al cuantizador es uniforme, o sea, si todos los posibles valores de las entradas son igualmente probables, es inmediato calcular el error promedio para diversos procedimientos de cuantización. Dividiendo en segmentos iguales de longitud unidad un único eje de coordenadas, y tomando como puntos de cuantización los centros de cada segmento, el error medio es 1/12, o sea, alrededor de 0,0833. Los mismos datos pueden ser bidimensionalmente cuantizados: para ello se los agrupa por pares; cada par de datos se considera un punto del plano. Dividiendo el plano en cuadrículas, y tomando por puntos cuantizadores los centros de las casillas, el error promedio es todavía de 1/12. Por otra parte, si el plano es dividido en hexágonos regulares de igual superficie que los cuadrados, y situando los puntos de cuantización en los centros de los hexágonos, el error medio de cuantización puede reducirse a $5\sqrt{3}/108$, aproximadamente, 0,0802.

Resulta muy llamativo que pueda mejorarse siempre la cifra de error, aun cuando los datos no estén distribuidos uniformemente. En su tesis doctoral, presentada en la Universidad de Stanford en 1963, P. L. Zador demostró que es siempre posible reducir el error promedio efectuando la cuantización en un espacio de dimensión superior. Es más eficiente esperar a que se hayan recogido varios puntos-dato y cuantizarlos luego todos ellos al mismo tiempo, considerándolos un único punto del espacio n-dimensional, que cuantizarlos uno por uno a lo largo de un eje. Al cuantizar, la pereza es rentable.

Desafortunadamente, el resultado de Zador, como antes el de Shannon, no es constructivo. El problema de hallar buenos cuantizadores multidimensionales sigue pendiente de solución, incluso con distribuciones uniformes de los datos. Hay, empero, varios empaquetamientos de esferas que producen, al parecer, excelentes criterios de cuantización. Fijémonos en el empaquetamiento de esferas en dos dimensiones, es decir, el empaquetamiento de cir-

cunferencias. Es sabido desde 1940 que las circunferencias quedan empaquetadas con máxima densidad disponiéndolas de modo que cada una se encuentre circundada por otras seis [*véase la figura 6*].

Imaginemos ahora que cada círculo del empaquetamiento esté envuelto por una membrana fina y elástica, y que se infla el interior del círculo. Al engordar los círculos, las membranas hacen presión unas contra otras, buscando rellenar los espacios libres del plano; si el hinchamiento es uniforme por todo él, cada círculo se expandirá hasta formar un hexágono regular. Como ya he mencionado, al basar en regiones hexagonales la cuantización de datos uniformemente distribuidos se minimiza el error promedio. Una expansión análoga de las esferas octodimensionales del empaquetamiento E_8 conduce también a un bajo error de cuantización, menor todavía que el error mínimo bidimensional. El problema general de cuantización, que requiere que el espacio sea descompuesto en, o recubierto por, una partición discreta, está cercanamente emparentado con el problema de hallar el recubrimiento esférico óptimo.



7. EMPAQUETAMIENTOS LAMINADOS de esferas *n*-dimensionales; se construyen adosando capas de un empaquetamiento reticular laminado de dimensión inmediatamente inferior, $n-1$. Por ejemplo, el empaquetamiento hexagonal bidimensional L_2 puede construirse apilando hileras de círculos cuyos centros estén situados según el empaquetamiento unidimensional Z^1. Análogamente, podemos apilar capa tras capa de esferas hexagonalmente empaquetadas, y producir así el empaquetamiento reticular L_3, el más denso de los conocidos en tres dimensiones. John Horton Conway, de la Universidad de Cambridge, y el autor han continuado la construcción; han hallado todos los retículos laminados en dimensiones menores o iguales que 25. Los retículos laminados en dimensiones 1 a 10, y en dimensiones 14 a 24, son únicos. En dimensión 11 hay dos retículos laminados, tres en cada una de las dimensiones 12 y 13, 23 en dimensión 25 y no menos de 75.000 en dimensión 26.

La búsqueda de empaquetamientos esféricos densos en espacios multidimensionales queda grandemente simplificada al fijar la atención en ciertos tipos de empaquetamientos llamados reticulares, cuyas configuraciones son sumamente regulares. Tomemos el empaquetamiento hexagonal de circunferencias recién descrito, y observemos que los centros de cualesquiera dos circunferencias adyacentes circundantes y el centro de la circunferencia núcleo forman un triángulo equilátero. Para calcular la densidad del empaquetamiento bastará determinar qué proporción de cada triángulo queda cubierta por círculos o partes de círculos. Dado que los triángulos teselan, o pavimentan, totalmente el plano, y puesto que la configuración determinada por los círculos dentro de cada triángulo es la misma, la densidad del empaquetamiento sobre un triángulo equilátero de los explicados es igual a la densidad de empaquetamiento sobre la totalidad del plano. Por geometría elemental, tal densidad es igual a $\pi\sqrt{3}/6$, o aproximadamente, 0,9069.

No hubiera sido posible el cálculo anterior de no haberse descubierto una unidad que, por repetición, llenase el plano. No obstante, es fácil imaginar que existen empaquetamientos de esferas sumamente irregulares, carentes de unidades repetitivas. Tales empaquetamientos son mucho más difíciles de estudiar que los repetitivos; no sólo es difícil o imposible calcular la densidad de los empaquetamientos irregulares, sino que también puede resultar imposible especificar las coordenadas de todos los centros. La propia definición de empaquetamiento reticular garantiza que puedan eludirse estas desventajas. Se dice que un empaquetamiento de esferas es reticular cuando, cada vez que contiene dos esferas, una de centro en el punto $(u_1, u_2, \ldots, u_n)$ y la otra con centro en $(v_1, v_2, \ldots, v_n)$, hay también esferas en el empaquetamiento centradas en todos los puntos $(a\,u_1 + b\,v_1, a\,u_2 + b\,v_2, \ldots, a\,u_n + b\,v_n)$, donde a y b son números enteros cualesquiera. Las coordenadas de los centros de éstas esferas se dicen engendradas o generadas a partir de las coordenadas de los centros de las primeras dos.

El más sencillo de los empaquetamientos reticulares es el reticular cúbico, en el cual las coordenadas del centro de cada esfera son números enteros; el retículo cúbico de número arbitrario n de dimensiones se denota por Z^n. El retículo "cúbico" unidimensional Z^1 está compuesto por segmentos rectilíneos discretos, todos de longitud unidad, centrados en los puntos de abscisa entera de la recta. Las "esferas", o segmentos rectilíneos, recubren el 100 por ciento de la recta; cada esfera es tangente a las otras dos. Por consiguiente, Z^1 resuelve los problemas de empaquetamiento y de número de osculación en una dimensión.

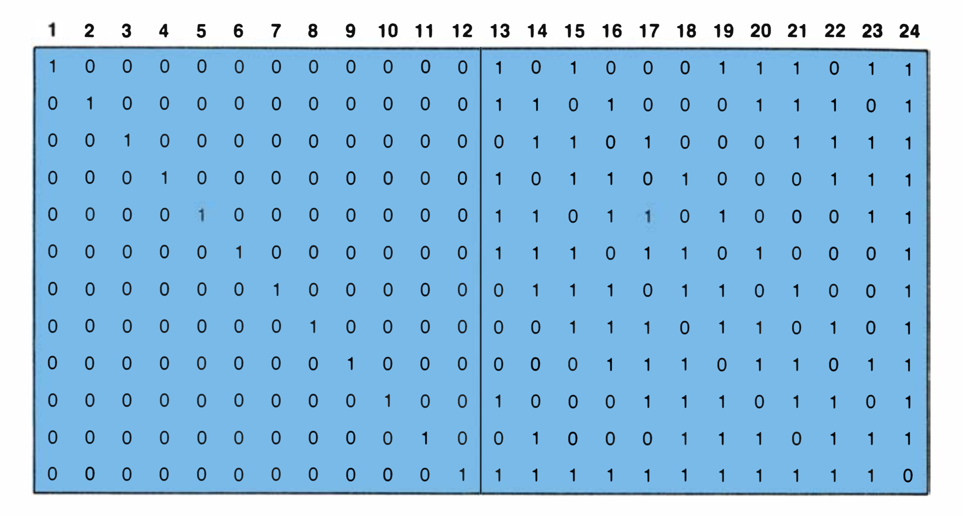
En dos dimensiones, sin embargo, el retículo Z^2 no es el más denso empaquetamiento. Su densidad de empaquetamiento es $\pi/4$, aproximadamente 0,7854, bastante inferior a la densidad de empaquetamiento hexagonal [*véase la figura 2*]. De la misma forma, la densidad del retículo Z^3 resulta bastante baja: es de $\pi/6$, alrededor de 0,5236. Puede generarse una familia mucho más densa de empaquetamientos reticulares centrando las esferas en puntos alternantes del retículo, a modo de cuadros de un damero. Para construir la nueva familia de retículos, coloréense alternativamente de rojo y negro los vértices de un retículo cúbico, y háganse coincidir los centros de las esferas con los puntos negros. De forma equivalente, los centros del nuevo empaquetamiento son los puntos de coordenadas enteras cuya suma sea número par. En dimensión arbitraria n, este empaquetamiento se designa por D_n. En D_3, por ejemplo, el origen (0, 0, 0) y el punto (1, 1, 0) son centros válidos de esferas, pero el punto (1, 0, 0) no lo es, porque $1 + 0 + 0$ es número impar.

La sucesión de empaquetamientos reticulares D_3, D_4, D_5, etcétera, reviste considerable importancia para el problema del empaquetamiento denso de esferas. El empaquetamiento D_3 es el retículo cúbico centrado en caras. Un modelo construido con bolas de pingpong muestra que una célula repetitiva del retículo es un cubo de dos unidades de lado, con una esfera en su centro; el radio de cada esfera es $\sqrt{2}/2$. La densidad del empaquetamiento puede calcularse hallando la porción del volumen del cubo que está ocupada por las esfe-

ras; resulta igual a $\sqrt{2}\,\pi/6$; alrededor de 0,7405. Aunque es posible que existan en tres dimensiones empaquetamientos esféricos más densos, Carl Friedrich Gauss demostró en 1831 que D_3 era el más denso de los empaquetamientos reticulares tridimensionales. Se sabe también que D_4 y D_5 poseen esta propiedad en espacios de dimensión 4 y 5, respectivamente.

A partir de la dimensión 5, sin embargo, D_n deja de ser el más denso de los empaquetamientos reticulares; para cuando se llega a D_8, quedan entre las esferas huecos muy grandes. Tan grandes son, que es posible infiltrar entre los huecos una segunda copia de D_8, sin solapamientos de las esferas. El retículo resultante se llama E_8. En 1934, H. F. Blichfeldt, de la Universidad de Stanford, demostró que E_8 era el más denso de los empaquetamientos reticulares octodimensionales y demostró también que ciertas secciones transversales de E_8, llamadas E_6 y E_7, constituían los empaquetamientos reticulares más densos de seis y siete dimensiones. En estas dimensiones no se han encontrado desde entonces empaquetamientos nuevos más densos, y de existir, tendrían que ser no reticulares.

En 1965, John Leech, por entonces en la Universidad de Glasgow, construyó un notable empaquetamiento en el espacio de 24 dimensiones; su construcción se esquematiza en la figura 8. El estudio del retículo de Leech, como suele llamarse a su empaquetamiento, ha llevado a una comprensión más profunda de las propiedades de otros retículos en dimensiones grandes, así como a importantes resultados en teoría de grupos. C. A. Rogers, razonando como hiciera para los empaquetamientos esféricos tridimensionales, dio cotas para las densidades máximas de empaquetamiento en espacios n-dimensionales cualesquiera. La acotación a respetar por todo empaquetamiento 24-dimensional es tan sólo ligeramente superior a la densidad del empaquetamiento de Leech. Cada esfera de este retículo contacta con otras 196.560. En 1979, A. M. Odlyzko, de los laboratorios Bell, y el autor demostraron que este número resuelve el problema de osculación en 24 dimensiones. El mismo método de demostración resolvió también el problema del número de osculación en ocho dimensiones. La respuesta es entonces de 240, que es el número de esferas tangentes a cada una en el retículo E_8. A estos dos resultados llegó también, por su cuenta, V. I. Levenshtein, del Instituto L. V. Keldysh de Matemática Aplicada, de Moscú. Incidentalmente, el problema permanece



1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	1	1	0	1	1
0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	1	1	1	0	1
0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	1	1	1	1
0	0	0	1	0	0	0	0	0	0	0	0	1	0	1	1	0	1	0	0	0	1	1	1
0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	1	1	0	1	0	0	0	1	1
0	0	0	0	0	1	0	0	0	0	0	0	1	1	1	0	1	1	0	1	0	0	0	1
0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	1	0	1	1	0	1	0	0	1
0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	1	0	1	1	0	1	0	1
0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	1	0	1	1	0	1	1
0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	1	1	0	1	1	0	1
0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	1	1	0	1	1	1
0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0

8. CONSTRUCCION DEL RETICULO DE LEECH, el más denso empaquetamiento de esferas de dimensión 24 hoy conocido. Se funda en las secuencias de 24 dígitos binarios aquí mostradas. El conjunto de todas las posibles sumas que puedan realizarse (en adición módulo 2) tomando los sumandos de entre estas doce sucesiones es un sistema de $2^{12} = 4096$ sucesiones binarias llamadas palabras de código. (Al sumar módulo 2, la suma de 1 y 1 es 0; el dígito de acarrero se desprecia.) Las 2^{12} palabras de código forman un código eficiente para transmisión de información, ideado por el norteamericano Marcel J. E. Golay en 1949. Los centros de las esferas del retículo de Leech son todos de la forma $2C+4X$ o de la forma $I+2C+4Y$, donde C es una palabra del código de Golay, I es el punto (1, 1, ..., 1) en 24 dimensiones, y X e Y recorren todos los puntos 24-dimensionales de coordenadas enteras. La suma de las coordenadas de cada punto X ha de ser número par; la suma de las coordenadas de cada punto Y, impar. Cada esfera tiene radio $2\sqrt{2}$; las más cercanas al origen poseen centros tales como ($\pm 4, \pm 4$, 0, 0, ..., 0), ($\pm 2, \pm 2, \pm 2, \pm 2, \pm 2, \pm 2, \pm 2, \pm 2$, 0, 0, ..., 0) y ($\pm 3, \pm 1, \pm 1, ..., \pm 1$). Cada esfera es tangente a otras 196.560.

abierto en todas las demás dimensiones, exceptuada las uno, dos y tres (en cuyos casos, la solución es de dos, seis y 12).

El retículo de Leech ha resultado indispensable a los especialistas en teoría de grupos para construir ciertos grupos finitos simples. Estos grupos son los bloques constituyentes de todos los grupos que constan de número finito de elementos; su clasificación, recién concluida, ha tenido ocupados a muchos matemáticos durante más de 50 años. En buena medida, los grupos simples desempeñan en teoría de grupos papel muy semejante al de los números primos en teoría de números, o los elementos químicos en química. Se han podido construir ciertos grupos simples importantes estudiando los conjuntos de todas las rotaciones rígidas y todas las simetrías del retículo de Leech que dejan fija la esfera central y permutan entre sí las esferas circundantes. El conjunto de estas operaciones recibe el nombre de grupo de simetrías del empaquetamiento; el grupo de simetrías correspondiente al empaquetamiento cúbico centrado en las caras en tres dimensiones puede verse en la figura 1.

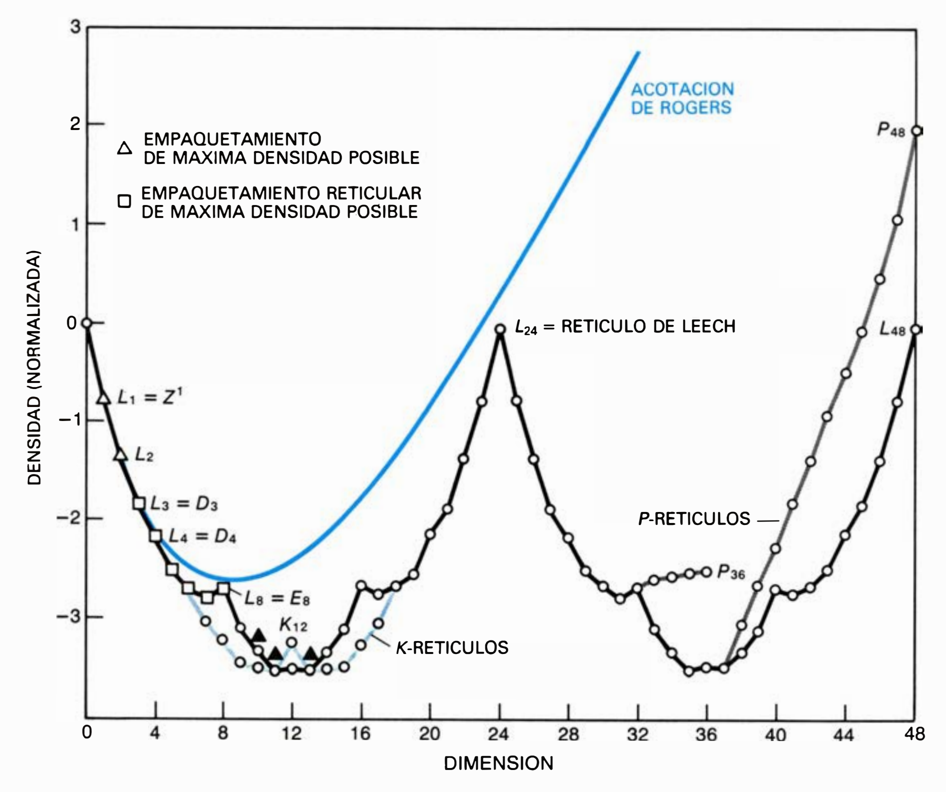
El grupo de simetrías del retículo de Leech fue determinado por John Horton Conway, de la Universidad de Cambridge, en 1968. Su orden, o sea, el número de elementos de que consta, es enorme, aunque no especialmente grande al situarlo en el contexto de la teoría de grupos: es $2^{22} \times 3^9 \times 5^4 \times 7^2 \times 11 \times 13 \times 23$, o sea, 8.315.-553.613.086.720.000. A partir de este grupo, que no es simple, J. Horton Conway construyó tres grupos simples anteriormente desconocidos, cuyos órdenes son en todos los casos divisores del orden del grupo de simetrías del retículo de Leech. En 1981, el retículo de Leech permitió a Robert L. Griess, Jr., de la Universidad de Michigan, construir uno de los últimos grupos finitos simples que se han encontrado. Bastante mayor que los grupos de Conway, ha sido apodado "el monstruo", pues el número de elementos de que consta es $2^{46} \times 3^{20} \times 5^9 \times 7^6 \times 11^2 \times 13^3 \times 17 \times 19 \times 23 \times 29 \times 31 \times 41 \times 47 \times 59 \times 71$, o sea, 808.017.-424.794.512.875.886.459.904. 961.710.-757.005.754.368.000.000.000. La construcción de Griess no es, empero, nada sencilla, y uno de los aspectos más fascinantes del retículo de Leech es que se tiene la impresión de que ha de existir una conexión más directa entre el grupo simple monstruo y él.

Tan denso resulta el empaquetamiento que produce el retículo de Leech que su influencia se deja sentir en todas las dimensiones inferiores. No es sorprendente que las rebanadas de buenos empaquetamientos produzcan buenos empaquetamientos en espacios de dimensión una unidad inferior. Por ejemplo, tajando D_3 queda a la vista una superficie de esferas hexagonalmente empaquetadas. Secciones transversales

apropiadas del retículo de Leech dan lugar a los más densos empaquetamientos conocidos en todas las dimensiones inferiores, a excepción de las 10, 11 y 13. Por ejemplo, una cierta rebanada octodimensional a través del retículo de Leech proporciona el retículo E_8.

Siendo posible construir empaquetamientos densos a partir del retículo de Leech, de arriba abajo, por así decirlo, es tentador plantearse cómo construir el retículo de Leech de abajo arriba, esto es, partiendo de empaquetamientos densos en dimensiones inferiores. Resulta que esta construcción puede llevarse a cabo de modo particularmente sencillo. Comencemos con el más denso de los empaquetamientos en una dimensión, el llamado Z^1. Con centro en el centro de cada esfera unidimensional de Z^1 se construye una esfera bidimensional cuyo radio sea ½. Ahora se construye otra capa de esferas bidimensionales idénticas a las primeras, y se la aloja en los huecos de la primera capa, lo más ajustadamente posible. Adosando de esta forma un número infinito de capas, el retículo bidimensional resultante es el empaquetamiento hexagonal denso; por su método de construcción, puede ser llamado "empaquetamiento laminar" en dos dimensiones, o brevemente, L_2.

He sugerido ya cómo pasar a tres dimensiones: se colocan bolas de radio ½ con centro en el centro de cada círculo de L_2, y se hacen encajar, en los huecos de la primera capa de bolas, nuevas capas de bolas idénticas, de modo que formen retículo. Dado que este empaquetamiento es equivalente a D_3, el procedimiento laminar da el mejor de los empaquetamientos conocidos tanto en dos como en tres dimensiones. Continuando de igual manera hacia dimensiones superiores, una dimensión por vez, los empaquetamientos reticulares resultantes son extraordinariamente densos. Desde hace mucho se sabe que L_4 y L_5 son equivalentes a D_4 y D_5, y que L_6, L_7 y L_8 son, respectivamente, equivalentes a E_6, E_7 y E_8. Así pues, los empaquetamientos laminares son los empaquetamientos reticulares de máxima densidad en hasta ocho dimensiones.



9. LOS MAS DENSOS EMPAQUETAMIENTOS conocidos de esferas de hasta 48 dimensiones han sido representados aquí según un método sugerido por John Leech. La densidad "normalizada" del empaquetamiento depende de la dimensión en que se efectúa. Su definición se funda en que la densidad del retículo 24-dimensional de Leech, dividida por el volumen de una esfera 24-dimensional de radio 1, es igual a 1. (El volumen de una esfera n-dimensional de radio 1 es igual a $\pi^{n/2}/(1\times2\times3\times\ldots\times n/2)$ cuando n es par y a $2(2\pi)^{(n-1)/2}/(1\times3\times5\times\ldots\times n)$ si n es impar). En un espacio de dimensión n arbitraria, este cociente se llama "densidad de centro", D. La densidad normalizada, mostrada en la gráfica, es igual a $\mathrm{Log}_2 D + n(24-n)/96$; para empaquetamientos reticulares laminados, la gráfica es simétrica respecto a la densidad normalizada del retículo de Leech. La ilustración muestra que L_3, L_8 y L_{24} se hallan muy próximos a la mínima de las cotas superiores conocidas para la densidad de los empaquetamientos de esferas. Los empaquetamientos reticulares laminados son los más densos conocidos en todas las dimensiones hasta la 32, si se exceptúan las dimensiones 10 y 13. Existe una segunda sucesión de retículos, llamada sucesión K_n, que arranca de L_6 y se reúne nuevamente con la sucesión laminar en L_{18}. Los empaquetamientos K_n son más densos que los L_n en las dimensiones 11, 12 y 13. Los empaquetamientos P_n también son reticulares; pero los empaquetamientos más densos de que se tenga noticia en 11, 12 y 13 dimensiones no son reticulares. Todos ellos han sido construidos a partir de códigos para la transmisión digital de información.

Recientemente, Conway y el autor han llevado un poco más lejos el procedimiento laminar: hemos hallado todos los retículos laminares en dimensión menor o igual que 25, y hemos construido ejemplos de tales retículos para cada dimensión, hasta 48. Resulta que a pesar de no existir más que un único retículo laminar para cada dimensión menor o igual que 10, hay dos procedimientos distintos de ir apilando capas del retículo L_{10}. Los dos retículos 11-dimensionales tienen la misma densidad, pero distintos números de osculación. Existen tres retículos laminares 12-dimensionales y otros tres 13-dimensionales, pero tan sólo uno de tales retículos para cada una de las dimensiones 14 a 24. El retículo laminar L_{24} es el retículo de Leech. Es únicamente en los espacios que tienen dimensiones 11, 12 y 13 donde se conocen secciones transversales del retículo de Leech superiores a las de los retículos laminares correspondientes.

Los retículos laminares se construyen capa por capa, haciendo encajar cada nueva capa tan ajustadamente como sea posible en los huecos de la capa precedente. Por tanto, la investigación de los retículos laminares está íntimamente relacionada con el problema de recubrimiento. Dilatemos cada esfera de cualquier empaquetamiento lo suficiente para englobar los huecos e intersticios; por definición, el empaquetamiento habrá de recubrir el espacio. En 1966, Leech conjeturó que si cada una de las esferas de su retículo fuese dilatada en un factor $\sqrt{2}$, las esferas recubrirían todos los puntos del espacio; tal recubrimiento pudiera muy bien ser óptimo en el espacio de 24-dimensiones. La conjetura de Leech era, no obstante, francamente difícil de probar, a causa, sobre todo, de la complejidad de las formas de los intersticios del retículo de Leech.

En el retículo bidimensional L_2 está claro que los huecos intersticiales del empaquetamiento son de una sola clase; la máxima distancia desde cualquier punto del retículo es siempre la correspondiente a un punto encerrado entre tres círculos; no hay, por tanto, nada que distinga a unos huecos de otros. Sin embargo, en el retículo de Leech, Conway, Richard A. Parker y yo mismo clasificamos 23 tipos distintos

de huecos antes de poder demostrar la conjetura de Leech. No sorprendentemente, resulta haber 23 formas distintas de apilar copias del retículo de Leech en el espacio 25-dimensional; hay pues 23 distintos retículos laminares 25-dimensionales. En dimensión 26, el número de retículos laminares es, cuando menos, 75.000.

Mucho menos se sabe de los espacios de dimensión mayor que 25. En 1905, Hermann Minkowski demostró que para cualquier dimensión, n, existen empaquetamientos reticulares de densidad mayor que 2^{-n}. Su razonamiento, como otros que he mencionado, no es constructivo. Para dimensiones pequeñas, la estimación de Minkowski es bastante burda; por ejemplo, para n igual a 24, enuncia que existen empaquetamientos cuya densidad supera unos 6×10^{-8}, cuando se sabe que el retículo de Leech es más de 32.000 veces más denso. En dirección contraria, Blichfeldt demostró en 1914 que, para valores de n arbitrariamente grandes, la densidad no podía superar un valor cercano a $2^{-0,5n}$. A pesar de grandes esfuerzos, no se obtuvo prácticamente mejora alguna sobre esta acotación hasta el decenio pasado, cuando G. A. Kabatiansky, Levenshtein y V. M. Sidel'nikov, en Moscú, demostraron que, en dimensiones n muy altas, la densidad no puede sobrepasar unos $2^{-0,599n}$.

E. S. Barnes, de la Universidad de Adelaide, A. Bos, de la N. V. Philips' Gloeilampenfabrieken, en Eindhoven, Holanda, Conway, Leech y yo hemos construido cierto número de empaquetamientos explícitamente dados en espacios de dimensión grande; pero ninguno de ellos tan denso como el teorema de Minkowski promete. Recientemente, Barnes y el autor han construido empaquetamientos reticulares, partiendo del retículo de Leech, hasta dimensión 100.000. La densidad de estos empaquetamientos es, aproximadamente, $2^{-1,25n}$ que a primera vista podría pensarse más o menos igual a la acotación del teorema de Minkowski. La verdad es que, en dimensión 65.536, nuestros empaquetamientos eran alrededor de $10^{40.000}$ veces más densos que ninguno de los empaquetamientos reticulares hasta entonces construidos. Pero, desafortunadamente, la forma exponencial de la acotación puede oscurecer el hecho de que aún estamos muy por debajo de la meta prometida. El teorema garantiza que quedan todavía por descubrir empaquetamientos 10^{4000} veces más densos.

Mineros y metalúrgicos antiguos en Huelva

Restos de explotaciones minero-metalúrgicas cubren parte de la provincia, confirmando así las referencias de los autores griegos y de la Biblia a las riquezas metálicas, sobre todo en plata, del mediodía de la Península Ibérica

Antonio Blanco Freijeiro

Es de justicia reconocer que uno de los tipos humanos que más ha contribuido a la formación del mundo civilizado es el herrero, el metalúrgico, el hombre cuyo tesón e inventiva proporcionó los metales a sus congéneres haciendo así posible el advenimiento del mundo moderno. Pero sus orígenes son muy remotos. Cyril Smith vislumbra así sus comienzos: "Uno de los primeros especialistas en salir de las sombras de la prehistoria es el herrero, un hombre que ya desde el principio ocupa un curioso puesto en la escala social. Unas veces se le ha tenido en gran estima, otras en poca, y esto último por razones obvias: aunque trabajaba mucho, era una persona sucia; su rostro solía estar tiznado de humo y sus ropas chamuscadas por el fuego. Pero, por otra parte, las cosas que hacía eran útiles y hermosas; poseía además la capacidad, que lo hacía asemejarse a los dioses, de saber alterar la naturaleza misma de la materia: podía convertir la roca opaca en reluciente metal y hacer de los materiales con que trabajaba algo líquido o sólido, rígido o flexible. En edades muy antiguas tales cambios eran contemplados como expresiones de fuerzas, a menudo espirituales, encerradas en la materia misma."

En la provincia de Huelva, conocida desde hace un siglo en el mundo de los mineros como "la California del cobre", las huellas de los más primitivos prospectores se encuentran apenas abandonamos los llanos de la franja costera y nos adentramos en las colinas de la sierra. Pisamos así el reborde meridional del que los geólogos llaman Cinturón Ibérico de Piritas. Allí, en las márgenes del río Corumbel, y en multitud de puntos de los términos de Valverde del Camino y de Zalamea la Real, asoman los primeros crestones de cuarzo y de calcita con los criaderos de cobre más fáciles de detectar y de explotar. Las hermosas vetas de color verdoso de la malaquita y azul de la azurita resaltan nítidas en las paredes y hendiduras de las rocas, vetas que a veces no alcanzan más que unos cuantos centímetros de ancho. Sería inconcebible que un fenómeno geológico tan llamativo no despertase la curiosidad del gran detector y coleccionista de piedras que fue el hombre prehistórico. Uno se imagina a uno de estos hombres extrayendo con su pico de pedernal un trozo de aquella sustancia quebradiza, de color tan atractivo, y pensando en qué provecho podría sacar de ella.

Quizá lo más obvio, lo primero, fuese utilizarla como colorante. Sabemos que los egipcios lo hicieron así, para maquillarse los ribetes de los ojos, y que esa operación revestía una gran importancia, incluso en el ritual religioso. Pero ignoramos si en la Huelva prehistórica se aprovechó el polvo de malaquita para esta o para otra finalidad. Nada tendría de raro que así fuese, pues no hay que olvidar que en los descubrimientos científicos no sólo interviene la curiosidad, sino muchas veces la sensibilidad de un artista. A este propósito recuerda Cyril Smith que los metales no fueron descubiertos porque algún individuo de la Edad de Piedra desease un material mejor para hacer herramientas. Este deseo sólo hubiera podido conducir al empleo de piedras mejores o de palos de formas más ingeniosas. "No –recalca este autor–, los metales hicieron su aparición porque hace varios milenios la sensibilidad artística de alguno fue excitada por una piedra interesante o bonita. Y aunque aquel impulso creador acabase luego dando lugar a una industria gigantesca, y aunque los metales hayan llegado a influenciar casi todas las esferas de la vida humana, los modos nuevos y sucesivos de trabajar con metales empezaron casi siempre en el dominio de las artes decorativas."

Pronto la malaquita dio pie a la obtención del cobre y éste a la fabricación de adornos (zarcillos, anillos, cuentas de collar, etc.) y de hachas y cuchillos más eficaces que sus predecesores de piedra. Para extraer la malaquita de los duros crestones de cuarzo en que se hallaba incrustrada, era menester romper éstos accionando unos pesados martillos de granito y de otras piedras duras, hechos de cantos rodados de forma más o menos ovalada, en cuya parte media los mineros hacían un surco anular que les permitía atarlos con una cuerda. La operación del martillado abría en la roca de cuarzo o calcita una trinchera irregular, como un foso de unos metros de profundidad. En las minas más antiguas se ve cómo la malaquita fue extraída en su mayor parte, pero no la calcopirita que por lo regular se encuentra debajo de la capa oxidada. En los alrededores de las trincheras no es difícil descubrir aún hoy, en la superficie del terreno, docenas de martillos de piedra del tipo descrito, rotos en su mayoría por el uso, pero siempre con el surco característico en su parte media.

Desde el año 1974 empezó a desenvolverse en la provincia de Huelva un

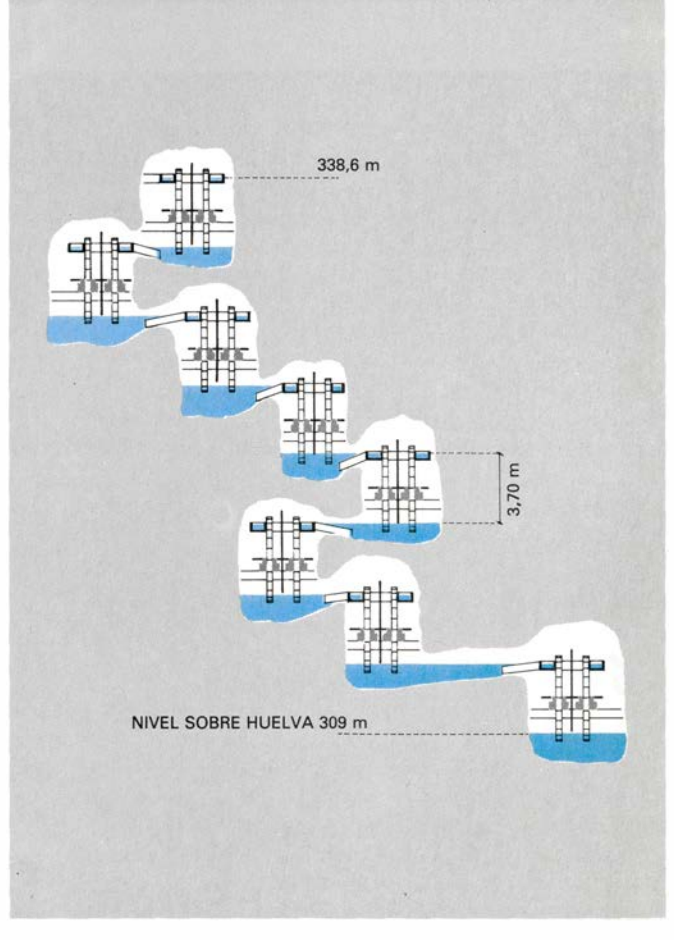
1. PLANTEAN LOS DOLMENES el problema de si existe una relación entre la metalurgía calcolítica y la primera manifestación arquitectónica que hizo uso de la piedra. En las cercanías de algunas minas calcolíticas se encuentran dólmenes saqueados como los que se divisan (*arriba*) desde el cerro de Chinflón. Abajo, restos de un dolmen de Pozuelo, construido de pedruscos hincados en el suelo y techado por otros colocados horizontalmente. En su tiempo no estaban descarnados como ahora, sino recubiertos de un gran túmulo de tierra. La identidad de la cerámica y el utillaje lítico apunta a la posibilidad de que los constructores de los dólmenes y los buscadores de metales fueran unos mismos.

proyecto de investigación arqueometalúrgica patrocinado por el IAMS (Institute of Archaeo-Metallurgical Studies, London) y financiado por las compañías de Río Tinto Zinc Ltd., de Londres, y las españolas de Unión Explosivos Río Tinto y Río Tinto Minera. Su objetivo primordial era estudiar la minería y la metalurgia antiguas de esta zona, cuya importancia había sido universalmente reconocida como uno de los focos primordiales de producción metalúrgica en el mundo antiguo. De esa importancia no sólo dan testimonio los autores griegos y romanos que ponderan el bronce y la plata de los tartesios, sino los millares de toneladas de escorias acumuladas por los romanos en lugares como Riotinto, Tharsis y Sotiel Coronada, sin contar las galerías y los ingenios mecánicos dejados *in situ* por los mismos agentes, para la extracción del mineral. El proyecto ha estado codirigido hasta ahora por Beno Rothenberg y por el autor de este artículo. Hasta el momento se han localizado y explorado unos 70 yacimientos en la provincia, y se han excavado un par de puntos clave localizados durante las exploraciones previas.

Cerca de la aldea de El Pozuelo, a unos 12 kilómetros al suroeste de Riotinto, una larga cadena de afloramientos de cuarzo atraviesa la campiña ondulada y hoy en día totalmente deshabitada. Hace poco, la repoblación forestal ha arado y sembrado de eucaliptos estas colinas, formadas por rocas volcánicas piroclásticas y ácidas, de modo que los crestones de cuarzo afloran como islotes desnudos y áridos entre la fronda de los plantones de eucaliptos. Al pie mismo de la colina donde se halla la mina de Chinflón, se encuentra un grupo de dólmenes. Estos monumentos plantean el problema de si existe una conexión entre la minería y la metalurgia calcolítica y esta manifestación arquitectónica que está hoy considerada como la primera en el mundo que hizo uso de la piedra. Los partidarios de la tesis de que los dólmenes pertenecen a una era plenamente neolítica negarán esta conexión: nosotros, que hemos encontrado en la propia mina de Chinflón útiles de piedra de tipo dolménico y cerámica de la misma filiación, tenemos por muy probable no sólo que los dólmenes sean de época calcolítica (Edad del Cobre), sino que éstos en particular pertenezcan a los mineros que por primera vez explotaron el yacimiento de Chinflón.

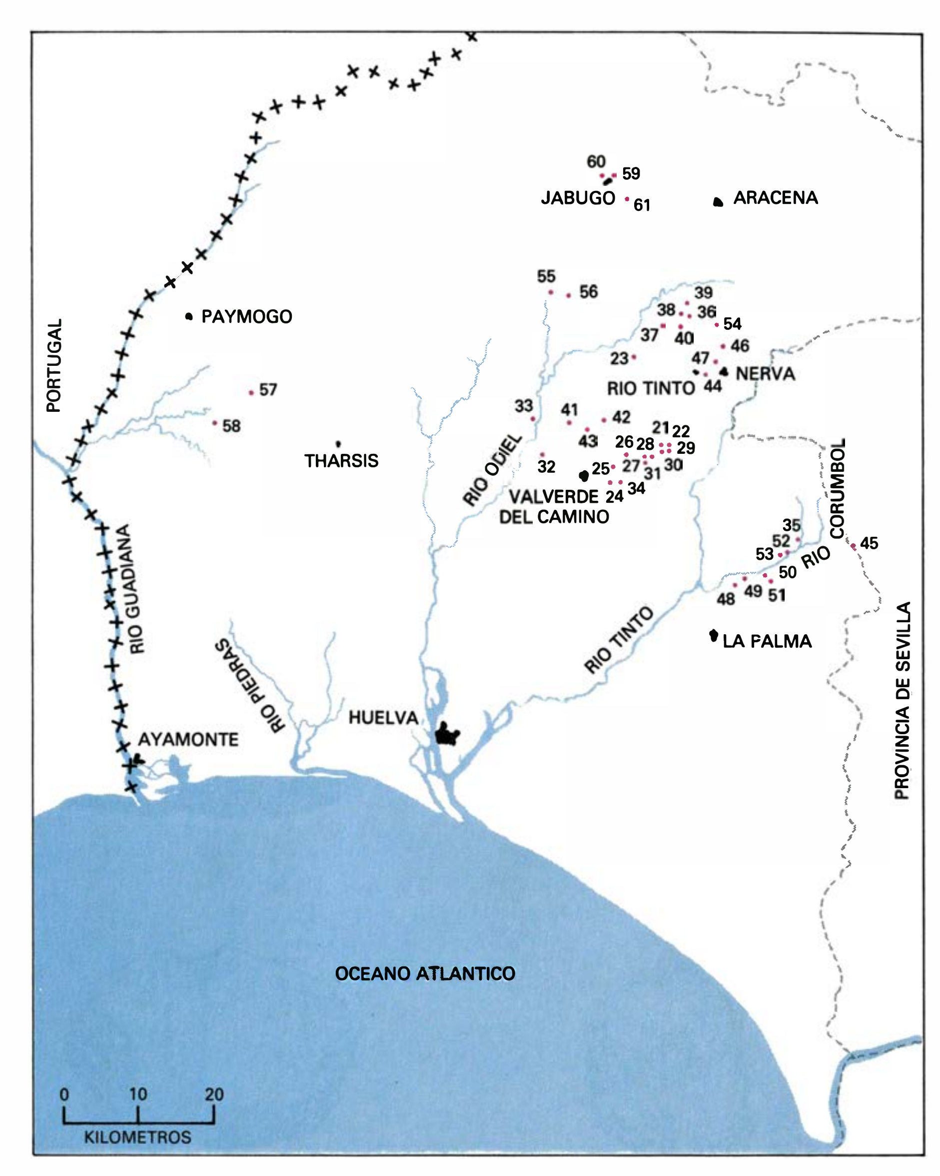
Cada uno de los cuatro dólmenes de Chinflón es, como todos sus congéneres, un enorme sepulcro colectivo en forma de galería, de unos dos metros de altura en su interior, construida a base de grandes losas monolíticas hincadas en el suelo y techadas con otras horizontales. Un túmulo de tierra, hoy desaparecido, cubría cada uno (o cada dos, cuando éstos se hallaban muy próximos) de esos monumentos. Hace poco Glyn Daniel ha tratado del tema en esta revista ["Monumentos megalíticos"; INVESTIGACIÓN Y CIENCIA, septiembre de 1980]. Los dólmenes de Chinflón habían sido expoliados varias veces, en tiempos antiguos y modernos, cuando Carlos Cerdán los dio a conocer en 1952 y rescató de ellos lo que los ladrones de tumbas habían desdeñado o no visto. Aun así, los ajuares dieron muestras de la riqueza que un día habían atesorado: muchos y finos objetos de piedra tallada y pulimentada; cuentas perforadas de serpentina, de fluorita, de calaíta y de otras sustancias raras y muy apreciadas; placas de pizarra de-





2. RUEDA ROMANA DE RIOTINTO, conservada íntegra, y esquema de la disposición en cadena de una serie de ellas (del Filón Sur) empleada cuando el desagüe de la mina no se podía hacer por una galería horizontal y era menester elevar el agua. Él giro de las ruedas había de hacerse con fuerza humana.

coradas con filas de triángulos rellenos de líneas incisas, perforadas en lo alto, y que se llevaban colgadas del cuello o prendidas del vestido; ídolos-placa de silueta antropomorfa; en suma, todos los elementos del mundo mágico-religioso del megalitismo andaluz, como los que nuestro Proyecto ha exhumado en un dolmen del término de Valverde que se hallaba en parecidas condiciones. Añádase a ello la cerámica y, en el dolmen n.º 4, un vestigio importantísimo: el trozo de un alfiler o punzón de cobre muy corroído. El cobre y los metales preciosos han sido el incentivo principal de los buscadores de tesoros, de modo que no debe extrañar que sólo un trozo de cobre de escaso valor se les haya escapado en este caso; pero para nosotros constituye un documento precioso por cuanto acredita que las personas sepultadas en aquellos dólmenes no sólo estaban en posesión de todo el utillaje propio del Neolítico, sino también de los productos de las primeras industrias del metal. Con ello no se demuestra –es cierto– que ellos beneficiasen la malaquita de la mina próxima, pero tampoco desdice de que lo hubieran hecho.



3. LA PROVINCIA DE HUELVA es objeto desde 1974 de un proyecto de investigación arqueometalúrgica patrocinada por el IAMS (Institute of Archaeo-Metallurgical Studies, London) y financiado por las compañías de Río Tinto Zinc Ltd., de Londres, y las españolas de Unión Explosivos Río Tinto y Río Tinto Minera. Su objetivo es el estudio de la minería y metalurgia antiguas de esta provincia. La investigación se ha verificado hasta ahora en las minas y yacimientos arqueológicos numerados en este mapa.

Ya he dicho cómo eran estas minas arcaicas. Ahora podemos comprobarlo: unas a modo de zanjas o trincheras de varios metros de longitud y de unos 80 centímetros de ancho, por término medio, abiertas en la masa de cuarzo, pero dejando aquí y allá, a ras del suelo, unas traviesas de la propia roca, para instalar en ellas los instrumentos –poleas, cuerdas, cestos– con que izar las cargas de mineral. La mayoría de las zanjas están ahora rellenas de tierra, pero su forma indica que los mineros primitivos se limitaban a seguir por el crestón la veta de mineral oxidado que les interesaba.

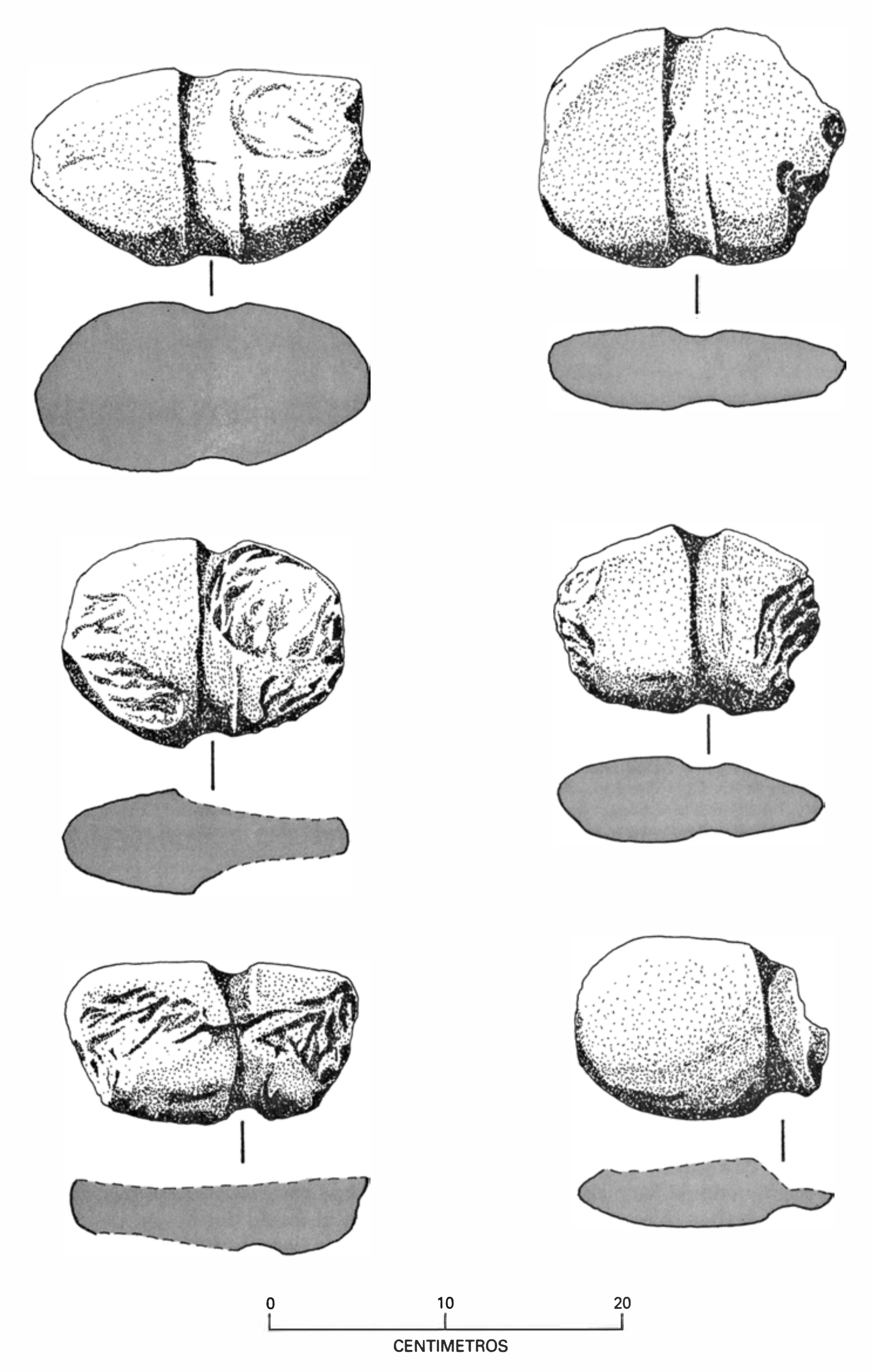
Las señales impresas en las paredes por las herramientas indican que sólo se empleaban en ellas los martillos romos y sin filo a que antes nos hemos referido. Muchos de ellos aparecen dentro y fuera de las zanjas. Las formas y los materiales varían, pero se trata siempre de piedras duras, volcánicas, que abundan como cantos rodados en los lechos de los ríos, donde sin duda los recogían antes de hacerles el surco de enmangue. Martillos de este género se encuentran en todas partes, desde Inglaterra hasta Arabia.

¿Qué ocurría una vez que el mineral había sido extraído por el procedimiento descrito? La única pista para averiguarlo son unos nódulos de escoria fayalítica, porosa, aislados y probablemente arrastrados desde el emplazamiento de un horno de fundición del tipo más primitivo que se conoce: un simple hoyo excavado en el suelo, rodeado de un murete de piedra y ventilado mediante un fuelle de piel de cabra o de otro animal. Tal vez en Masegoso hubiésemos dado con uno de esos hornos si los tractores de la repoblación forestal no se nos hubiesen adelantado. Lo único que encontramos fue el fondo del hoyo y una de las piedras hincadas que lo rodeaban. Hoy por hoy sólo tenemos hornos bastante completos en el Sinaí, y además con las mismas escorias que en Huelva. En esos hornos la malaquita era mezclada con un mineral de hierro y sometida a la acción del fuego que liberaba al cobre. El hoyo no tenía ni toberas (boquillas de barro que protegen el extremo del fuelle) ni revestimiento de paredes. Dado lo rudimentario del esquema, la escoria no alcanzaba en él un estado de fluidez que permitiese extraerla por un agujero de salida, sino que quedaba aprisionada en el horno. Cuando éste se apagaba y se enfriaba, era preciso extraer la escoria y luego desmenuzarla para sacar de ella las diminutas pepitas de cobre que se habían formado en su interior. Es posible que en el lento proceso de enfriamiento una parte de este cobre se depositase en el fondo del horno en forma de régulo o lingote. El procedimiento, modesto y laborioso, no podía más que alcanzar mínimos resultados: a lo sumo, unas docenas de kilogramos al término de cada campaña, pero de un metal tan precioso entonces como el oro y la plata. Por eso sus dueños podían llamarse ricos.

El tipo de horno descrito y deducible del tipo de escoria producida en él permitía alcanzar la temperatura de 1180-1350 grados Celsius, necesarios para la fundición de la mezcla. Ello implica el uso de fuelles, aunque hasta ahora no se hayan encontrado señales de éstos ni en Huelva ni el Sinaí.

Por rudimentario que parezca el sistema, su empleo comportaba el conocimiento de los principios prácticos de la fundición del cobre, aunque no, naturalmente, el de la química correspondiente. No podemos por menos de admirar a estos metalúrgicos de la Huelva de hace unos 5000 años, conscientes de las ventajas del hoyo en el suelo, cuyas paredes habían de recrecer mediante un murete de piedras que rodeasen la boca. Aunque no llegasen a él, se aproximaban al horno de cúpula que a lo largo de la historia se ha venido empleando en muchos países de la Tierra. Sabían estos metalúrgicos, además, cómo preparar el mineral y hacer con él una carga bastante equilibrada para el horno. Sólo una larga experiencia en esta operación podía garantizar el éxito de la misma, como ha comprobado en la práctica R. Tylecote.

No es de creer, sin embargo, que éste fuese el sistema de fundición más antiguo empleado por el hombre, aunque sí lo sea para nuestros conocimientos actuales de la historia de la tecnología. Pero dejemos a un lado las especulaciones y atengámonos a los hechos. Si una metalurgia de este tipo existía ya en la Huelva del III o incluso tal vez del IV milenio a. C., ¿es que fue inventada allí o introducida de otra parte, como podría ser, pese a la distancia, el Cercano Oriente, donde está acreditada su existencia por aquellas fechas?



4. GRANDES MARTILLOS Y PICOS de granito y de otras piedras duras eran utilizados por los mineros primitivos para extraer la malaquita (óxido de cobre) de los crestones de cuarzo o de calcita en que se hallaba incrustada. Se aprovechaban para estas herramientas pesados cantos rodados, recogidos en los cauces de los ríos y dotados por el hombre de un surco anular que permitía atarlos a un palo con una cuerda. Muchos se rompían en cada jornada de trabajo, pero no era difícil ni caro reemplazarlos.

En otros yacimientos de Huelva, como el de Cuchillares, al norte de Riotinto, y ahora ya de lleno dentro de la cuenca del Cinturón de Piritas, nuestras exploraciones han localizado venas de malaquita y de azurita entre filetes de "gossan" (óxidos y sulfatos de hierro) y en la superficie de afloramientos de toba y de riolita volcánica. Si aquellos óxidos y carbonatos de cobre se hubiesen echado al horno en compañía del gossan en que están empaquetados, el resultado hubiera sido el de una fundición. En Cuchillares se dan, por tanto, unas condiciones óptimas para que un hombre de ingenio hubiese ideado la metalurgia del cobre. Muchas mazas de piedra de tipo muy rudimentario y algunas escorias como las antes descritas fueron recogidas por nosotros en aquel lugar. Las pruebas de que los mineros primitivos lo habían frecuentado eran, pues, irrebatibles.

El panorama arqueológico que de aquella época ofrece la provincia de Huelva hace pensar que la sociedad prehistórica alcanzó en ella el estadio de la civilización calcolítica en virtud de su propia dinámica, sin necesidad de explicarla recurriendo a la tesis difusionista y menos aún a una invasión de colonizadores. La riqueza de los ajuares de los dólmenes y la metalurgia de Cuchillares-Chinflón pueden explicarse como fruto de las energías de la sociedad neolítica hispana sin necesidad de agentes ni de estímulos foráneos. Llegó un momento en que el cobre, sin ser excesivo, pudo ser objeto de un lucrativo comercio y éste convertirse en moderada fuente de riqueza para aquella sociedad primitiva.

Una de las varias sorpresas que nuestra exploración nos ha producido es la

de ver cómo esa floración cultural e industrial carece de continuidad en el Calcolítico reciente y en la Edad del Bronce antigua. El vaso campaniforme, que sería de esperar en el primero de estos períodos, y la cerámica y los bronces argáricos característicos del segundo, no están presentes en la zona minera, o por lo menos nosotros no hemos advertido su presencia. La falta del campaniforme podría tener su explicación en el hecho de que las minas más antiguas de Huelva, las minas de malaquita, de forma de trinchera, existentes en los afloramientos de cuarzo, se cerraron durante la fase última del Calcolítico, cuando estaba en uso el vaso campaniforme y en que ni entonces ni más tarde –la primera Edad de Bronce– la provincia de Huelva acertase a descubrir un modo de suplir aquella deficiencia. El hecho es que, por lo que ahora sabemos, en Huelva no se fabricaron los bronces arsenicales propios de aquellas épocas, ni bronces de ninguna otra clase antes de llegar a los bronces de estaño; mas para eso hay que alcanzar la etapa del Bronce Final, también llamado Bronce Atlántico. La conclusión a que por ahora hemos llegado es que los hombres del Calcolítico acabaron con las minas de malaquita beneficiables, llegaron a ver y tocar las piritas que había debajo, pero las dejaron por no saber qué hacer con ellas.

Pero la verdadera riqueza minera de Huelva no estaba en los afloramientos de óxidos de cobre como la malaquita y la azurita, ni siquiera en el cobre nativo que ocasionalmente pudiera aparecer; la riqueza se hallaba a mayor profundidad, fuera del alcance de las herramientas y de la tecnología del hombre calcolítico, en los macizos de piritas con sus zonas de enriquecimiento secundario de cobre, plata y oro situadas por debajo de las monteras de gossan y que pueden alcanzar un espesor de 100 metros. Estas monteras llegan a los 30 metros de profundidad. Su dureza y resistencia son tales que todos los intentos de perforarlas sin las herramientas metálicas adecuadas estaban condenados al fracaso. Como no se encontrase, pues, una vía de acceso, como podría ser una cueva que las aguas hubieran abierto en la ladera o al pie de una montaña, estos criaderos eran inaccesibles para el hombre prehistórico; de ahí que, salvo contadísimas excepciones, ninguno de ellos fuese explotado antes de comienzos de la Edad de Hierro.



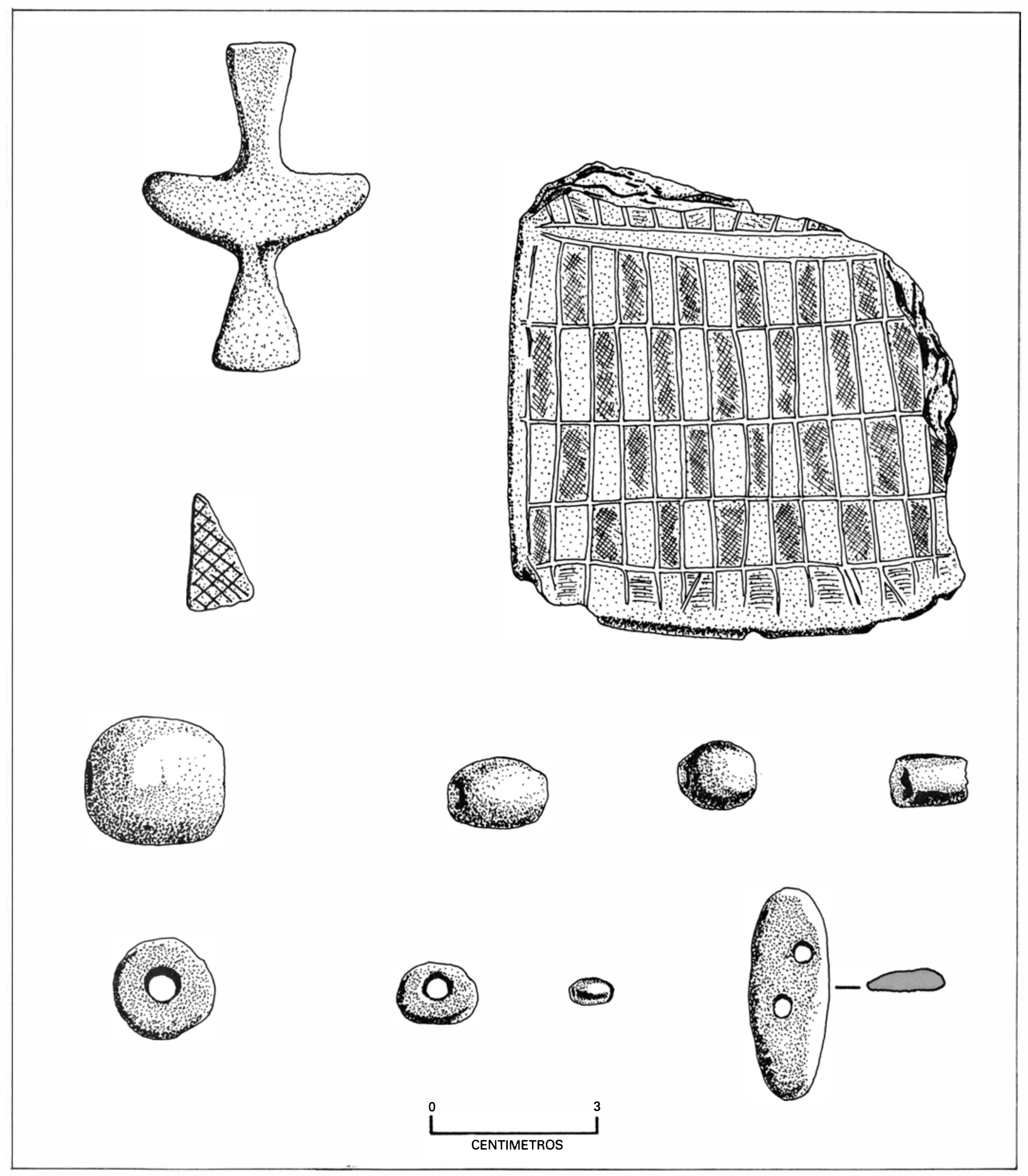
5. TRINCHERAS IRREGULARES como esta de Chinflón eran abiertas a martillazos en los afloramientos de roca siguiendo las vetas de malaquita y azurita. En las minas más antiguas se aprecia cómo fue extraída la totalidad aprovechable de estos minerales; pero no así la calcopirita, que siguió intacta.

En la época del Bronce Final (*circa* 1200-800 a. C.) se reanudan las actividades mineras en Huelva con dos sensibles novedades tecnológicas: el empleo de herramientas de metal para horadar la roca (sus huellas se advierten en las menudas entalladuras que dejan en las paredes) y los hornos revestidos de barro y ventilados a través de las toberas en que se enchufaban los fuelles. Los picos y cinceles, primero de bronce y más tarde de hierro, no sólo permitirán trazar galerías más regulares que las mazas o martillos de piedra, sino horadar rocas como el gossan, cuya dureza hacía inoperantes a los útiles de piedra. Nace así un nuevo tipo de mina, que en Chinflón comienza donde los mineros calcolíticos se habían detenido: el pozo circular, de unos 80 centímetros de diámetro (lo imprescindible para que un hombre pueda descender), con peldaños en la pared de la roca y galerías horizontales de extracción, circulación y desagüe. Los datos obtenidos por nuestra exploración indican que este tipo de mina, que para abreviar llamaremos "de pozo y galería", comienza a ser implantado durante el Bronce Final, bastante antes de que los fenicios hiciesen su aparición. Ha de considerarse, por tanto, una innovación autóctona de los tartesios o de quienesquiera que fuesen los portadores de la cerámica bruñida. Lo lógico y lo probable sería que entonces cayesen en desuso los martillos de piedra de la minería calcolítica, pero es posible que por rutina o por inercia algunos operarios siguiesen empleándolos. Eran, sin embargo, herramientas obsoletas, inservibles para abrir los nuevos tipos de galería.

Agotados ya por los hombres del Calcolítico los veneros de malaquita, la nueva época se concentró en la zona de enriquecimiento secundario, cubierta por las monteras de gossan. En Riotinto el descubrimiento y el primer acceso a éstas pudo producirse en unas cuevas hoy desaparecidas, pero que se conservaron hasta el siglo pasado: la Cueva del Tabaco y la Cueva del Lago, donde nacía el río Tinto. Las operaciones de extracción no tardaron en descubrir que además de los minerales de cobre había otros dos mucho más rentables en la nueva época: los de la plata y el oro. Téngase en cuenta que pese a la enorme importancia de la producción de cobre, todavía hoy Riotinto produce plata (25 toneladas en 1979) y oro (2,5 toneladas en 1979). La plata va a ser desde ahora el primer producto de las minas de Huelva. Habrá incluso épocas como la romana republicana en que la producción de cobre quede completamente desatendida. De ahí que las grandes concentraciones de escoria de las principales minas romanas y actuales (Riotinto, Tharsis, Sotiel Coronada) sean casi por completo escorias de plata. Por consiguiente, la riqueza argentífera de Tartessos de que las fuentes históricas se hacen eco no es una fantasía ni una leyenda, sino una reali-

dad perfectamente comprobable en los análisis de laboratorio.

En lugar del cobre puro y del arsenical de las etapas precedentes, el metal de la nueva industria es un bronce de estaño en su gran mayoría, el "bronce tartésico" del que dice Pausanias que había un armario en el tesoro de los sicionios en Olimpia. Las herramientas y las armas no se forjan, sino que se funden en moldes bivalvos; de éstos salen casi completas, sin necesidad de ulterior elaboración en frío. Los metalúrgicos no sólo producen una gran variedad de armas y objetos de bronce, sino recipientes y joyas de plata y oro, de una tecnología cada vez más compleja y refinada. La metalurgia extractiva es igualmente mucho más sabia: los hornos emplean fundentes de óxido de hierro perfectamente controlados; la escoria es una verdadera fayalita, derretida en delgadas placas que permiten romperla en trozos pequeños y extraer así todo el cobre que hayan retenido. Lingotes en forma de torta, y de tres a cuatro kilos de peso, eran los régulos que



6. PIEZAS DEL AJUAR funerario de un dolmen de Los Gabrieles, cerca de Valverde, con muchos y finos objetos de piedra pulimentada: un ídolo cruciforme entero, fragmentos de dos ídolos-placas y varias cuentas perforadas. La mayor de éstas es de fluorita, piedra que no existe en Huelva y hubo de ser importada por lo menos de la sierra de Córdoba; las otras son de calaíta, piedra que sólo conocemos a través de la arqueología y de serpentina.

se depositaban en el fondo del horno de aquella época.

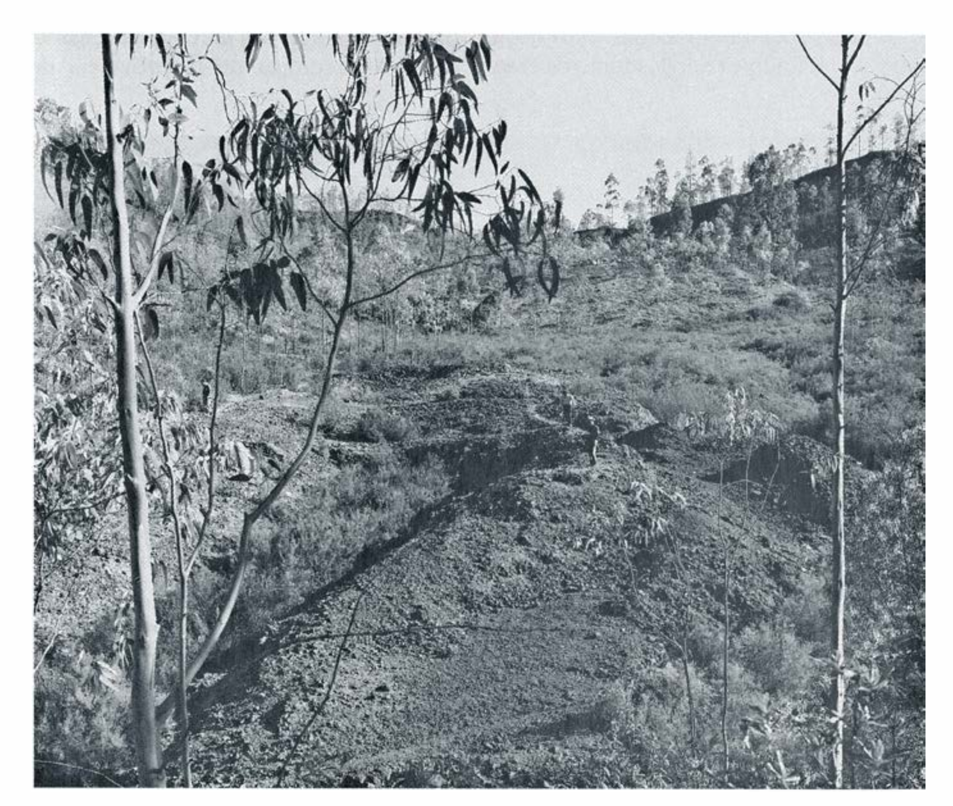
Durante toda la duración de ésta, y por sorprendente que parezca, Riotinto no produce cobre, sino exclusivamente plata. Ignoramos si en las otras grandes minas (Tharsis y Sotiel) ocurre lo mismo. Aún no hemos logrado explicarnos el por qué de este exclusivismo. El cobre tenía que seguir siendo muy valioso. Las numerosas armas y objetos de bronce dragados en el fondo de la ría de Huelva acreditan el intenso tráfico comercial a que estaban sometidos los productos de aquella industria. Y sin embargo, los mineros de Riotinto no se ocupaban de ella. Tal vez entonces tuviesen más a mano los minerales de plata que los de cobre. A veces ocurre en las minas que la logística de la producción se ve condicionada por la disponibilidad de los minerales. Es posible que las excavaciones den otra solución al problema. Lo que hasta ahora hemos comprobado, por lo que a la zona de Riotinto se refiere, es que entre las escorias antiguas de cobre, que se hallan amontonadas en las márgenes del Tinto, fuera del área especializada en la plata, no aparece nada anterior a la época romana imperial.

Otro resultado sorprendente de la exploración (aunque tal vez sólo sorprendente al primer pronto) es que la venida y el establecimiento de los fenicios en Cádiz y otros puntos de la costa no comportasen novedad alguna en el terreno de la tecnología y de la industria; que las cosas siguiesen tal y como estaban hasta entonces, con su sello occidental, atlántico. Ni siquiera sabemos si la introducción del hierro, cuando ésta se produce, fue debida a ellos. Si, como es de suponer, llegaron a la Península pertrechados de armas y herramientas de hierro, esa novedad no tuvo repercusión en la zona minera, muy rica en minerales ferruginosos, hasta por lo menos el siglo VII a. C. Hasta entonces las labores realizadas en las minas siguen adoleciendo de la misma tosquedad que antaño. La situación no mejorará sensiblemente hasta época romana, en la que las galerías están impecablemente labradas.

En síntesis, pues, hay que reconocer que los fenicios "orientalizaron" la cultura tartésica, aportando a ella multitud de elementos suntuarios –los artículos de prestigio y los signos externos de una civilización superior–, provocando con su gran demanda de metales preciosos e industriales el aumento de la productividad de las minas, pero dejando que éstas siguiesen en manos de quienes hasta entonces las habían bene-



7. VARIOS SISTEMAS DE EXPLOTACION MINERA suelen coincidir en un mismo yacimiento como en este de Cueva del Monje: trincheras calcolíticas, abiertas con martillos de piedra; pozos y galerías labradas con cinceles de bronce desde la época de este metal, y de hierro más tarde, trataron de economizar al máximo las labores de perforación, y por tanto, terriblemente angostas. En el campo de la técnica minera los fenicios no aportaron novedad alguna respecto a lo que sabían los aborígenes tartesios.



8. TONELADAS DE ESCORIAS romanas como éstas de Sotiel Coronada se acumulan en los vertederos y en los cauces de muchos arroyos de las minas de Huelva explotadas en tiempos del imperio de Augusto y de sus sucesores. Fue una empresa de titanes encaminada a la obtención de la plata y el oro de las zonas de enriquecimiento secundario, pero sin desdeñar el cobre y el hierro que también proporcionaban las grandes minas. Con todo, el producto principal de minas como Riotinto fue la plata. Eliminada la competencia de las empresas privadas de época republicana, la explotación constituía un monopolio estatal, al que estaban supeditadas todas las actividades de la provincia minera, sometida a una explotación exhaustiva.

ficiado, los tartesios y los célticos, establecidos de antiguo en la zona y dueños de una tradición minera y metalúrgica.

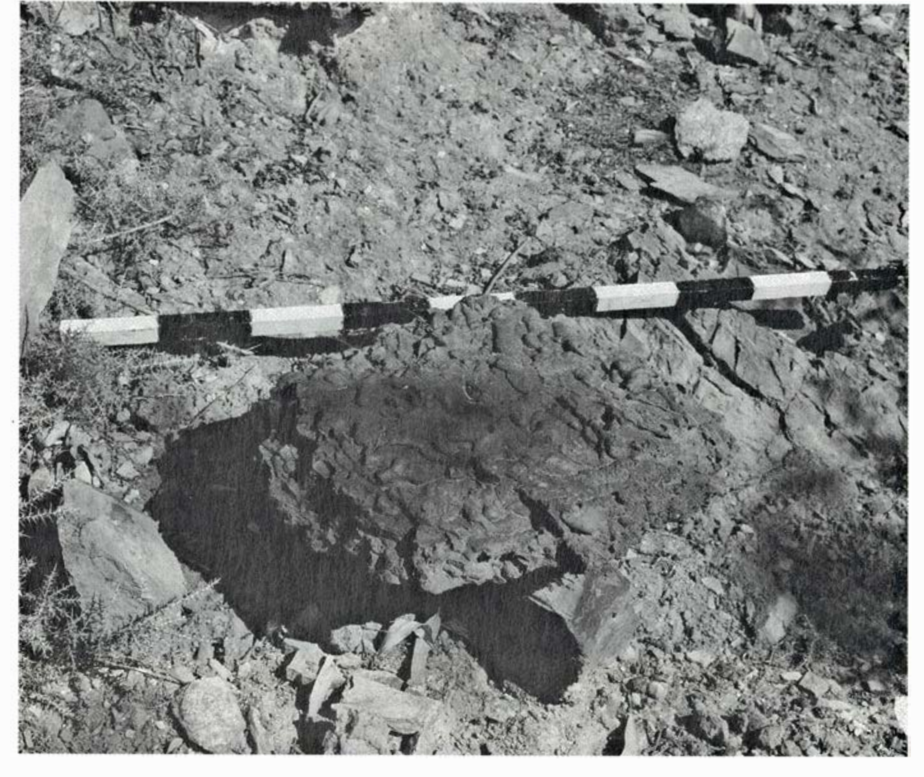
La tecnología de la fundición de plata que estos pueblos hispanos habían desarrollado en tiempos del Bronce Final permanece inalterada durante la etapa de la colonización fenicia. Las escorias contienen fragmentos de la cerámica incisa y digitada y de la cerámica bruñida propias del mundo indígena. De vez en cuando aparece con ellas algún tiesto de una anforilla oriental: es el primer síntoma de que la conexión fenicia se ha producido ya. Pero la tecnología sigue siendo la misma, la propia de los tartesios. En una fase posterior a la de los primeros contactos, la cerámica experimenta un cambio: gran parte de ella (cerca del 70 por ciento) está hecha a torno, no sólo la vajilla de mesa, sino la de cocina, y, por supuesto, también las ánforas del aceite, del vino y de las salsas y conservas de pescado. Esta situación se mantiene hasta principios del siglo v a. C. La uniformidad y homogeneidad de las capas de escoria, verdaderamente impresionante a lo largo de todos estos siglos, demuestra que nada cambia en los sistemas de fundición, que no se produce una colonización industrial en los dominios de las minas y que la metalurgia indígena persiste inmutable.

¿Ha de sorprender el que esto haya sido así? Sabido es que los fenicios eran un pueblo de marinos y mercaderes, naturales de la costa de lo que hoy es el Líbano: eran también diestros artesanos en la manufactura de textiles, de vidrio, de marfil, de bronce y de objetos de metales preciosos. La Biblia recuerda que Salomón les confió la construcción y la decoración del Templo de Jerusalén. Pero por carecer de minas en su país de origen, los fenicios carecían de conocimientos y de experiencia en el campo de la metalurgia extractiva, especialidad que en el mundo antiguo requería la maestría y la experiencia de muchas generaciones. Esto último era justamente lo que poseían los tartesios, y por ello no debe sorprender que los fenicios no interviniesen en la labor de éstos. No cabe duda de que sus colonias y factorías realizaron una labor comercial y cultural que alteró profundamente la fisonomía del país y a la que vinieron a sumarse, al mismo tiempo o después, los griegos y los cartagineses; pero con todo y con eso, la industria metalúrgica siguió siendo, según todas las apariencias y por lo que los autores antiguos manifiestan, una empresa de los indígenas.

Las convulsiones que la Península Ibérica experimenta en el siglo III a. C. con la conquista de gran parte de ella por Cartago primero y por Roma después, no dejaron una huella sensible en la arqueología de la provincia de Huelva, ni dentro de las minas ni fuera de ellas. Es posible que la propiedad de los yacimientos pasase a los nuevos dominadores y que la mano de obra experimentase también profundos cambios, considerando el auge adquirido por la esclavitud en el mundo romano con posterioridad a las Guerras Púnicas; pero estas son cosas que la arqueología difícilmente registra, y menos aún la mera exploración de superficie realizada por nosotros en la mayor parte de los casos. La única excavación extensa y detenida –la de Tejada la Vieja– realizada durante la prospección nos reveló que en un momento dado, probablemente a principios del siglo IV a. C., aquella población quedó desierta. El abandono definitivo y el traslado de sus habitantes a otro lugar no muy distante, pero situado en el llano y no en la cima de una colina, parece más la consecuencia de una orden de arriba que de una decisión propia y espontánea, pero no sabemos de quién partió la orden en cuestión. Sospechamos del mando cartaginés, pero nada más.

Al producirse, dentro ya de la época romana, el cambio de régimen de la república al imperio, el panorama de las minas cambia por completo. En todas partes queda bien de manifiesto que el estado, y no la empresa privada, ha asumido la dirección y la administración de las explotaciones. La compra en unos casos y la expropiación forzosa en otros fueron los procedimientos empleados. La historia y hasta la toponimia (en el nombre de los Montes Marianos) conservan el recuerdo de Sexto Mario, propietario de una mina de la sierra de Córdoba, residente en Roma, cuyos bienes fueron confiscados por Tiberio, y él mismo arrojado de la roca Tarpeya, en virtud de una acusación calumniosa. Ignoramos si en Huelva se cometieron abusos de poder como el acabado de reseñar, pero las consecuencias fueron las mismas.

La política minera de la administración imperial acarreó dos consecuencias inmediatas: la apertura de nuevas minas y la ampliación de las hasta entonces existentes. En unas y otras fue implantado con carácter general y con el sello inconfundible de las obras bien hechas, propio de Roma en todos los órdenes de la ingeniería, el sistema de pozos y galerías. A la producción de plata se suma la de cobre, pero no entrando en las masas de calcopirita, sino manteniéndose en la zona de enriquecimiento secundario. Las concentraciones de pirita cuya parte superior oxidada había sido beneficiada por los mineros calcolíticos y después la explotación continuada a mayor profundidad por



9. CONO DE ESCORIA solidificada, de medio metro de diámetro, resultante de varias fundiciones sucesivas en un mismo horno de extracción de plata. Un hombre del oficio tiene aún hoy mucho que admirar en una pieza como esta, por la maestría técnica que acredita y por la perfección del trabajo.

los de la Edad del Bronce Final permanecieron intactas en época romana. Así la mina de Chinflón ha conocido tres etapas de labores, separadas por pausas multiseculares: Calcolítico, Edad de Bronce, siglos XIX-XX. Es evidente por ello la falta de interés de los romanos en ese tipo de mineralización que no se volvió a sentir hasta la edad contemporánea. En la misma Riotinto, donde los romanos tanto trajinaron, no hay señales de extracción de calcopirita. Las galerías que parecían penetrar en la masa de ésta, y que al principio nos desorientaron, iban en realidad encaminadas a la muy rica calcocita, apreciadísima entonces y ahora.

Las inversiones de energía, de técnica y de capital fueron inmensas. Salvo las zonas agrícolas de Villalba del Alcor y de la Palma del Condado donde vemos instalados a los capitalistas agrarios, toda la provincia de Huelva concentra su actividad en la producción y en la exportación de metal –cobre, oro, plata y hierro–. Ni en el siglo XIX ni hoy se trabajó en una mina que no hubiera sido explotada ya por los romanos. La epigrafía nos informa ahora de la variedad de gentes ocupadas en las minas, en su mayoría forasteras: germanos, cántabros, emeritenses, olisiponenses (de Olisipo = Lisboa), lusitanos de Talábriga (al sur de Oporto), galaicos de la cuenca del Limia, y otros. Los grilletes que se encuentran ocasionalmente en las galerías confirman la presencia en ellas de condenados a trabajos forzados, los *damnati ad metalla*, otra prueba de que es el estado el explotador de la mina, pues una empresa particular no hubiese podido emplear esta fuerza laboral. Con éstos alternan los trabajadores libres y los esclavos.



10. PANORAMA DE UNA GRAN MINA romana, dedicada en Sotiel Coronada a la extracción de plata. En primer término, una bocamina cuadrada, hecha con el esmero y la competencia característicos de las obras de Roma. Lo mismo todo el sistema de galerías, desagües e ingenios de extracción de agua.

Aun nivel intermedio, por encima de esta mano de obra, mercenaria y carcelaria, y por debajo del *procurator metallorum* y de su equipo de funcionarios, se hallan los técnicos que Roma puede permitirse traer del Oriente o de cualquier otra región del Imperio con tal de que sea personal altamente cualificado como geólogos, prospectores, constructores de hornos, fundidores y lo que hoy llamaríamos ingenieros industriales, que podían ser esclavos o libertos imperiales. De esta clase solía ser el *procurator metallorum*, el hombre de confianza del emperador. Así Pudens, liberto del emperador Nerva, detentó la autoridad máxima en Riotinto en tiempos de Trajano.

Con toda la región minera concentrada en la explotación de sus minerales, la producción de cada uno de sus centros no podía permitirse fallos ni demoras. La eficacia del sistema se pone de manifiesto en el volumen y en la calidad de las escorias: perfectas escorias de derretido, fondos de horno con las múltiples capas de sucesivas fundiciones, que no requerían ningún trabajo en frío realizado con posterioridad. En todas partes se trabaja bien, sin descanso, sin escatimar los medios ni el esfuerzo. En las grandes minas se van acumulando las escorias por miles de toneladas, dando testimonio de la multitud de obreros empleada en ellas y de la eficiencia de los responsables de la organización del trabajo, de la disciplina y de los técnicos puestos al servicio de la gran empresa estatal.

Estas grandes concentraciones de masas laborales en parajes remotos y poco acogedores, con dificultades de vivienda, de provisión de alimentos e incluso de agua para beber, debió de exigir una vasta red de abastecimientos y transportes que sólo el estado se hallaba en condiciones de proporcionar. Por eso, aunque las leyes hablan de sociedades, y tales sociedades están acreditadas en la península (verbigracia la *Societas Castulonensis*, de la zona de Linares, pero presente también en Lora del Río, Sevilla), el apogeo de tales empresas debió de tener lugar en época republicana y tal vez a comienzos de la imperial. Al menos por ahora, y puesta nuestra mirada preferentemente en Huelva, no podemos por menos de suscribir la afirmación de Rostovtzeff: "la tendencia gradual de la política imperial en materia de minas y canteras consistió en eliminar paulatinamente a los grandes capitalistas y concentrar la explotación de aquéllas en manos de los funcionarios del estado...".

Uno de los muchos problemas a resolver por la Exploración será el del combustible empleado por las minas de época imperial. Si todas las combustiones de los hornos se hubiesen realizado a base de carbón de leña –único tipo de carbón empleado en la Antigüedad– la totalidad de los encinares de la Península Ibérica no hubiera bastado para abastecer las minas de Huelva. Hasta tal punto es esto cierto, que un técnico como L. Salkield no ha tenido reparo en suponer, con las cifras de Riotinto y del suroeste en la mano, que la pobreza maderera de España arranca de la época romana y de las operaciones mineras realizadas en aquel entonces. Nosotros tenemos en estudio otra posibilidad, sin descartar el empleo archiprobado del carbón de leña: la de que junto con éste se hiciese uso de otro combustible derivado de las piritas. De no haber sido así, podríamos estar seguros, en efecto, de que cuando las minas redujeron su febril actividad a fines del siglo II d. C. no quedaba una encina en su radio inmediato de acción y muy pocas en el área de hasta donde podían llegar sus proveedores.

Juegos de ordenador

Altibajos de los números pedrisco. A la búsqueda del algoritmo general

Brian Hayes

Tres pasitos adelante, dos pasitos atrás. Aunque no es la forma de viaje más eficiente, parece seguro que acabaremos por llegar. Un curioso problema de teoría de números, aún no resuelto, pone en duda tal conclusión. El problema puede enunciarse como sigue. Tomemos un número entero positivo cualquiera, que llamaremos N. Si el número N elegido es impar, lo multiplicamos por 3 y le sumamos 1 al producto; es decir, reemplazamos N por $3N + 1$. Si el número N ha sido par, lo dividimos por 2, sustituyendo N por $N/2$. En ambos casos, el número obtenido se toma como nuevo valor de N, y se repite el procedimiento. Tras muchas iteraciones, ¿tenderán los números a ir haciéndose cada vez más grandes, o cada vez más pequeños? ¿Convergen hacia algún valor concreto, o divergen hacia el infinito? ¿Cuánto se tardará en dejar sentada la suerte que habrá de correr el número?

Dado un valor concreto de N, la respuesta a estas preguntas apenas si requiere algo más que aritmética elemental. Por ejemplo, si N es 27, número impar, el valor siguiente es $(3 \times 27) + 1$, o sea, 82; a continuación vienen 41 y 124. Habrá, claro está, muchos altibajos en esta serie de números; su valor crecerá cada vez que N sea impar, y descenderá cuando N sea par. Se invita al lector a prolongar la sucesión y ver adónde conduce.

La dificultad no estriba en calcular la sucesión correspondiente a un valor concreto N, cualquiera, sino en dar una solución general, válida para todo posible valor de N. Hasta la fecha no se ha conseguido. Se han ensayado muchos números, y en todos los casos la conducta se ha manifestado conforme a una misma pauta; pero nadie ha conseguido demostrar que todos los números habrán de acomodarse a ella. Mal podrá decirse que se trate de uno de los problemas más importantes de la teoría de números, pero sí que es uno de los más irritantes. El procedimiento es fácil de describir y llevar a cabo, pero resulta sorprendentemente difícil comprender lo que realmente sucede.

El problema ilustra perfectamente la utilidad y las limitaciones del ordenador digital en tanto que instrumento de investigación matemática. Para explorar más allá de los enteros pequeños es imprescindible alguna clase de auxilio mecánico que se encargue de efectuar los cálculos; prácticamente todos los ordenadores servirán, incluso las calculadoras de bolsillo programables. Por otra parte, para ampliar el cálculo a una gama de enteros significativamente mayor, se precisan las más potentes maquinarias de cálculo. Y cuando se abordan las cuestiones verdaderamente profundas, es dudoso que ningún ordenador sirva de nada. En su mayor parte, el ordenador es una herramienta de "matemática experimental": sirve para generar ejemplos y contraejemplos. Descubrir principios generales en los vaivenes y peregrinaciones de N exige teoremas, no trituradoras de números.

Al aplicar repetidamente la regla de transformación a un número arbitrario, ¿qué resultado cabe esperar? He aquí tres hipótesis ingenuas:

El primer razonamiento reza así: Hay igual cantidad de enteros pares e impares; por tanto, en cualquier serie larga de cálculos deben presentarse por igual valores pares e impares de N. Cuando N sea impar, quedará aumentado en tres veces su valor (y un poco más); en cambio, cuando N sea par, quedará solamente dividido por 2. Por consiguiente, tras muchas iteraciones, el valor de N habrá de crecer ilimitadamente. Por término medio, en cada iteración su valor debería quedar multiplicado por $(3N + 1)/2$, que para valores grandes de N es casi exactamente $3N/2$.

La segunda hipótesis se funda en que todo lo que sube tiene más tarde que bajar. El razonamiento comienza con la observación de que en cuanto el cálculo produce una potencia exacta de 2, la serie de números sufre una cascada de divisiones consecutivas por 2, cayendo en picado hasta el valor 1. (Exceptuado el propio 2, al dividir por 2 cualquier potencia de 2 resulta necesariamente un número par, con lo que se elige invariablemente la rama descendente del cálculo.) Hay entre los números naturales una infinidad de potencias de 2, por lo que, llevando el cálculo todo lo lejos que sea necesario, acabaremos por toparnos con una de ellas. En el transcurso del cálculo pudieran alcanzarse valores muy elevados de N: más dura será la caída.

Aunque en su forma el tercer razonamiento es parecido al segundo, la conclusión a que conduce es diferente. Observemos que cada vez que el cálculo cambia de dirección, como acontece al encontrar un término impar tras una serie de términos pares, vuelve, por así decirlo, a territorios donde ya estuvo antes. En efecto, en su vagabundeo arriba y abajo de la recta numérica, puede retornar a un dominio finito de números, con frecuencia arbitraria. Es de esperar que al cabo se tropiece con algún valor ya visitado, y en cuanto así acontezca, el futuro del cálculo queda fijado definitivamente. Puesto que el procedimiento para elegir el paso siguiente es plenamente determinístico, tan pronto se repita un valor entraremos en un bucle a ejecutar indefinidamente.

No hay que tomarse demasiado en serio las tres hipótesis recién presentadas. No pueden ser las tres correctas. Algunas de sus premisas pueden, lícitamente, ponerse en tela de juicio. En particular, todas se inspiran en consideraciones probabilísticas; pero las sucesiones de números engendradas según la regla nada tienen de aleatorias. ¿Qué puede decir al respecto la experimentación matemática?

El lugar donde comenzar la experiencia es al principio, tomando $N = 1$. Se trata de un número impar, y las ins-

trucciones ordenan multiplicarlo por 3 y sumarle 1. El resultado, 4, es par; hay por tanto que dividirlo por 2, produciendo otro número par; una nueva división entre 2 nos devuelve al punto de partida, 1. Por tanto, el primer ensayo otorga una bonita confirmación a dos de las teorías especulativas antes citadas. Tal como predice la teoría de caída en cascada, el cálculo se topa con una potencia de 2, y lo hace así justamente en la primera iteración. Del mismo modo, tal como predice la teoría cíclica, el cálculo queda atrapado en un bucle sin fin, repitiéndose indefinidamente los valores 4, 2 y 1.

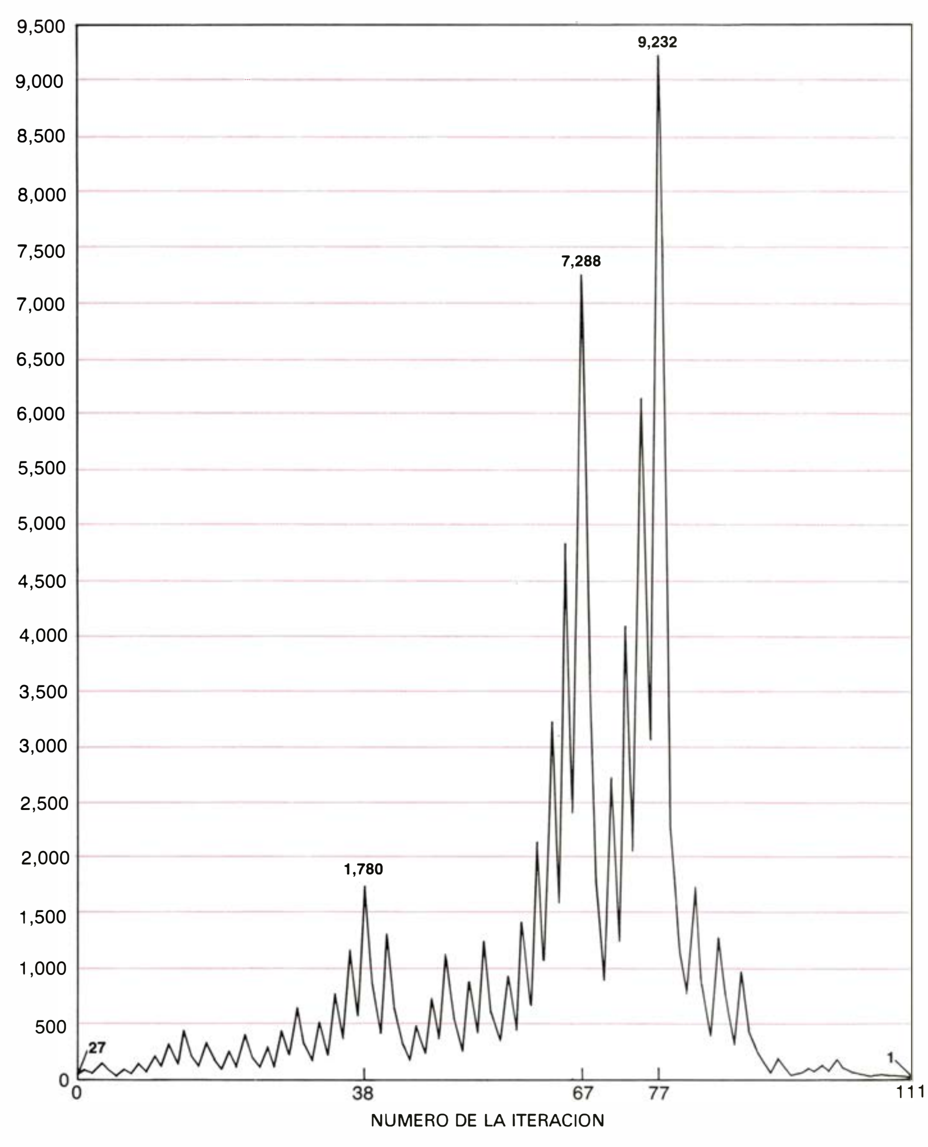
Entre los números naturales, el 1 es muy especial: es el primero y el mínimo. Consiguientemente, los resultados obtenidos en el caso $N=1$ pudieran ser atípicos. Antes de llegar a conclusión ninguna deberíamos hacer algunas comprobaciones más. Puesto que la suerte de 2 y de 4 es ya conocida por el estudio del caso $N = 1$, el candidato natural para el nuevo ensayo es 3. Es impar, y así el valor siguiente será $(3\times3) + 1$, o sea, 10. Al dividir por 2, resulta 5, y después, tras multiplicar por 3 y sumar 1, el resultado es 16. Nuevamente hemos obtenido una potencia de 2, y la sucesión se desploma, pasando por $N = 8$ y entrando en el bucle 4-2-1.

Examinados los cuatro primeros números naturales, la tendencia parece manifiesta. Hay, empero, motivo para ponerla en duda. En los cálculos realizados hasta ahora hay dos cantidades destacables: el máximo de los valores alcanzados por N en el proceso del cálculo y la longitud de la trayectoria por él seguida, que definiré como número total de iteraciones necesarias para alcanzar el valor 1. Para el propio 1, el valor máximo es 1, y la longitud del camino, cero. Para $N = 2$, el valor de pico es 2, y la longitud, 1. Para 3, el máximo alcanzado es 16, y la longitud del camino, 7. El ejemplo $N = 3$ sugiere que el valor máximo alcanzado y la longitud de la serie pueden ser mucho mayores que el valor inicial de N, y que, por ello, puede que la sucesión resulte no acotada para ciertos valores de N.

Fijémonos ahora en la sucesión que se engendra cuando el valor inicial es 27. Como ya se hizo notar, los tres primeros números son 82, 41 y 124; después, dos divisiones sucesivas retrotraen la sucesión a 31. Así que después de cinco pasos, apenas si hemos avanzado. Empero, conforme prosigue el cálculo, el mecanismo de "tres pasitos adelante, dos pasitos atrás" va suscitando una serie de oscilaciones de amplitudes cada vez mayores. Se alcanzan nuevos picos en 142, 214, 322 y 484. Hay también fuertes retrocesos (en el paso 19 el valor ha caído hasta 91), pero la tendencia continúa ascendente. El cálculo pasa por 700, por 1186 y por 2158, y en la septuagesimaséptima iteración el valor alcanzado ya es importante: 9232. Tal parece que ya hemos cogido vuelo. Resulta, sin embargo, que el camino termina en 1 tras un total de 111 pasos, sin jamás remontarse por encima de 9232. (Puede verse la trayectoria completa en la figura 1.)

Se han efectuado cálculos del tipo que acabo de explicar para todos los enteros de una gama extraordinariamente amplia. Nabuo Yoneda, de la Universidad de Tokio, ha ensayado todos los valores hasta 2^{40}, aproximadamente $1{,}2 \times 10^{12}$. En todos los casos el resultado ha sido el mismo: tras un número finito de altibajos, la serie queda confinada en el bucle 4-2-1, donde habrá de permanecer por siempre. De entre los primeros 50 enteros, el 27 es el que más larga ruta sigue para regresar al 1 (si bien las trayectorias del 31 y 41 no son mucho más cortas, y alcanzan el mismo valor de pico, por razones obvias a la vista de lo ya explicado). No se ha descubierto ningún entero positivo que genere una serie capaz de remontar el vuelo hacia el infinito; tampoco se han descubierto bucles distintos del 4-2-1. Por otra parte, la conjetura de que todos los números se ajustan a esa misma pauta sigue desprovista de fundamentación teórica.

El "problema $3N + 1$", como acostumbra a ser llamado, tiene un pasado nebuloso, pero no parece ser de gran antigüedad. A lo largo de los últimos 30 años ha rebrotado repetidamen-



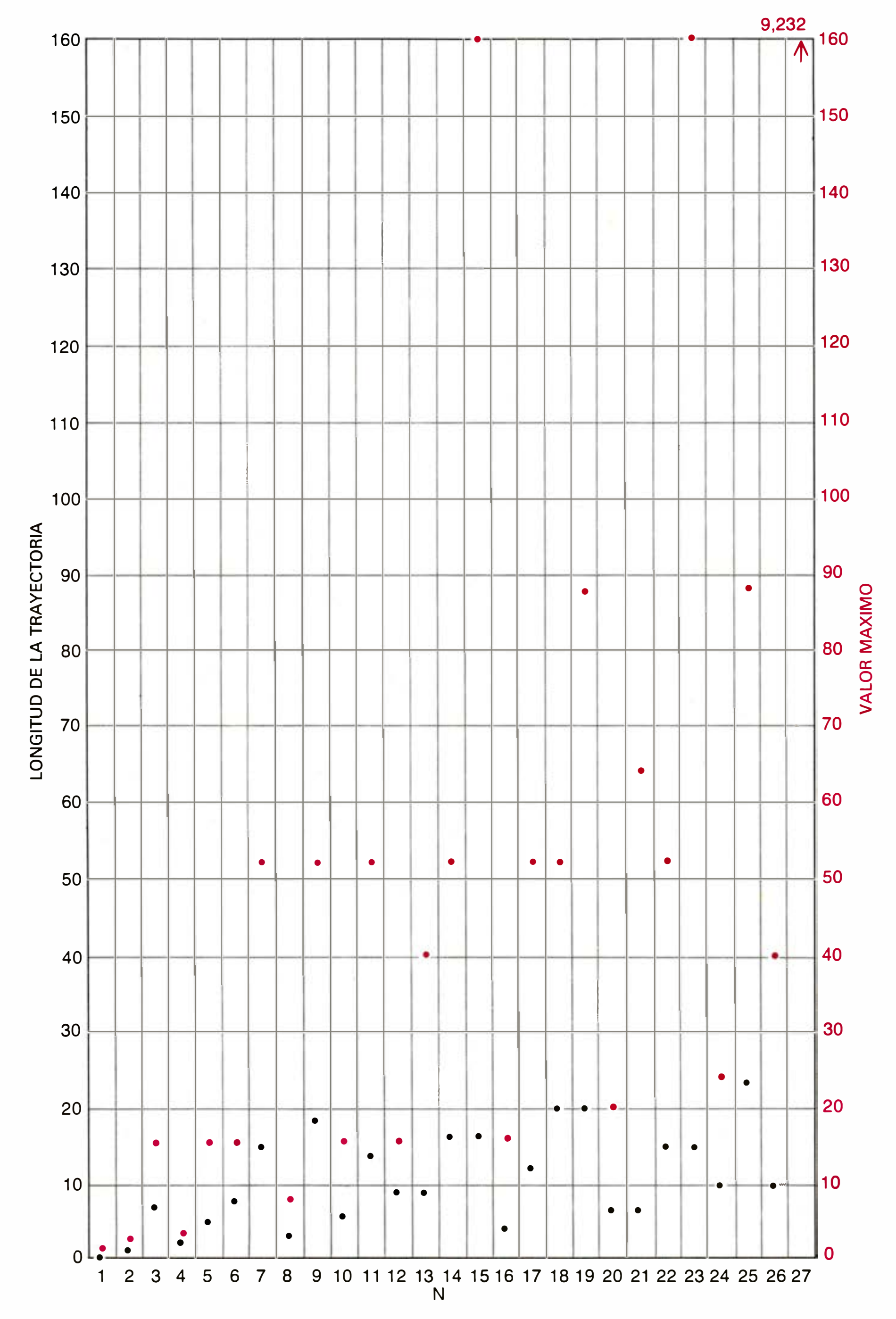
1. Sucesión de números pedrisco que comienza en 27

te en los departamentos de matemáticas y de ciencias de cómputo de diversas universidades, con idas y venidas de apariencia tan caprichosa como las de la propia sucesión. Jeffrey C. Lagarias, de los Laboratorios Bell, quien recientemente ha estado husmeando acerca de los orígenes del problema, y estudiando las expectativas que hay de resolverlo, hace observar que el problema parece haberse inventado varias veces. En el decenio de 1930, Lothar Collatz, a la sazón alumno de la Universidad de Hamburgo, investigó una clase de problemas entre los que se cuenta el "3N + 1", si bien su trabajo no se publicó hasta muchos años después. En 1952, e independientemente, el matemático inglés B. Thwaites redescubrió el problema, y años más tarde fue inventado de nuevo por Richard Vernon Andree, de la Universidad de Oklahoma en Norman.

Lagarias cita unos veinte artículos de investigación del problema "3N + 1" y de sus generalizaciones, casi todos ellos publicados a lo largo de los 10 últimos años, pero el problema había circulado de boca en boca desde mucho antes. Un colega de Collatz, Helmut Hasse, lo llevó a la Universidad de Syracuse hace unos 30 años; desde allí, Stanislaw Ulam lo trasladó a Los Alamos, y de allí, quién sabe adónde. Shizuo Kakutani, quien tuvo noticia del problema hacia 1960, comentó a Lagarias: "Durante un mes, todo Yale estuvo trabajando en él, sin resultado. Fenómeno similar se produjo al mencionar yo el problema en la Universidad de Chicago. Incluso llegó a decirse, en broma, que el problema formaba parte de una conjura para entorpecer la investigación matemática en los Estados Unidos".

A principios del decenio pasado, un grupo integrado en el Laboratorio de Inteligencia Artificial del Instituto de Tecnología de Massachusetts emprendió un ataque vigoroso y sostenido contra el problema, cargando el acento, sobre todo, en el cálculo numérico con auxilio de ordenador. El problema está censado como Item 133 de las transacciones (informales e inéditas) del grupo, llamadas HAKMEM.

En sus idas y venidas, el problema ha recibido multitud de nombres distintos. Llamarlo "problema 3N + 1" no parece plenamente satisfactorio, pues carga el acento en una de las mitades del procedimiento, pasando por alto la otra. De las diversas posibilidades, aquélla con la que más congenio califica a los números engendrados por el procedimiento a partir de un valor inicial de "números-granizo". La senda que va recorriendo la sucesión recuerda mucho la trayectoria de las bolitas de pedrisco en el seno de las nubes de tormenta, que suben impelidas por las corrientes ascendentes, pero acaban cayendo por su propio peso.



2. Longitud de trayectoria (**negro**) *y valor máximo alcanzado* (**rojo**) *para cada uno de los primeros 27 enteros*

Con lenguajes de alto nivel, como BASIC, se pueden escribir programas de ordenador que calculen números-granizo en unas pocas líneas. Más todavía, el núcleo del algoritmo puede quedar plasmado con tan sólo un enunciado, que en BASIC podría adoptar la forma

```
IF N MOD 2 = 0 THEN N  = N/2
ELSE N = 3*N + 1
```

Las personas podemos efectuar la primera operación (a saber, determinar si el número N es par o non), sin cálculos explícitos, pero no así los ordenadores. N MOD 2 es una operación por congruencia, que calcula el resto de la división de N entre 2. Si el resto es 0, se ejecuta la parte THEN del enunciado,

y se le asigna a N el valor $N/2$. De no ser así, se ejecuta la parte ELSE, y N toma el valor $3N + 1$.

Los programas en BASIC funcionan bastante bien para engendrar "pedriscos" mientras el número inicial no exceda de unos pocos centenares; para emprender exploraciones más ambiciosas, estos programas son de lentitud intolerable. El enunciado en BASIC requiere una división (que forma parte de la operación "módulo") una comparación, y después, ya una segunda división, ya una multiplicación seguida de una adición. Las operaciones de multiplicación y división requieren mucho tiempo, sobre todo en ordenadores pequeños. Hay en este caso mucho que ganar hablándole a la unidad de proceso directamente en su lenguaje nativo. De esta forma podemos prescindir de las operaciones de división y multiplicación.

En la figura 5 vemos, esquematizada, la descripción de uno de estos programas en lenguaje de máquina. Se supone que el valor de N se encuentra inicialmente en un registro llamado AX, que sirve también como "acumulador", en el cual se efectúan las operaciones aritméticas. El valor alojado en él, al comenzar el procedimiento, es la representación binaria de 27.

El primer paso consiste en guardar una copia del valor inicial en otro registro, aquí etiquetado BX. La operación de dividir puede evitarse sacando partido de una propiedad del sistema binario de numeración: correr hacia la derecha las cifras de un número expresado en el sistema binario equivale a dividirlo por 2, lo mismo que al correr un lugar hacia la derecha las cifras de un número decimal, éste queda dividido por 10. En el proceso de corrimiento, la cifra situada en el extremo derecho (la de unidades) queda guardada en una célula de almacenamiento con capacidad para un bit, llamada "señalizador de acarreo" (*carry flag*, en inglés). El contenido de este señalizador determina si el número era par o non, pues en notación binaria todos los números impares terminan en 1 y, todos los pares, en 0.

Si N es par, el cálculo acaba aquí, pues el valor que permanece en el registro AX tras el desplazamiento de un lugar a la derecha es el cociente $N/2$. En este caso, sin embargo, N es impar, y se requieren más operaciones. En primer lugar, recobrar el primitivo valor de N, que permanecía alojado en el registro BX. Después, en lugar de multiplicarlo por 3, este valor se suma consigo mismo dos veces. Aunque procediendo así se requieren dos instrucciones de máquina en lugar de una, su ejecución es considerablemente más rápida. El paso final es incrementar en 1 el número contenido en AX. Con el juego de instrucciones de un determinado procesador, la secuencia completa requiere 20 cuentas de reloj del ordenador cuando N es par y 18, cuando N es non. A frecuencias de reloj que rondan los 5 megahertz, el fragmento de programa podría, en teoría, ejecutarse unas 250.000 veces por segundo. (Podrían todavía economizarse unos cuantos ciclos de reloj, a costa de cierta pérdida de claridad en el programa.) El algoritmo equivalente, empleando operaciones de multiplicación y división, exige 175 ciclos cuando N es par y 286 ciclos de reloj cuando N es impar.

Los registros representados en la ilustración disponen de ocho bits, y por tanto pueden acomodar números menores que $2^8 = 256$. En casi todos los microprocesadores, los registros son en realidad de 16 bits de anchura, lo que permite manejar números menores que 65.536. Incluso este límite supone una restricción muy seria; un programa que utilizase aritmética de 16 bits sería incapaz de calcular la "granizada" de valores iniciales mayores que 702. Para conseguir mayor capacidad hace falta aritmética de múltiple precisión, en la cual un mismo y único número es escindido entre dos o más registros o localizaciones de memoria. Con precisión de 32 bits pueden representarse números de hasta unos cuatro mil millones; 64 bits llevan el límite hasta 10^{19}. Empero, toda mejora en precisión está gravada por pérdida de velocidad.

El algoritmo de cálculo de un valor de N es tan sólo un fragmento de un programa capaz de funcionar. Tiene que haber, además, medios que permitan introducir valores en la máquina y la presentación de resultados. Un conjunto de programas encaminados a explorar "pedriscos" tendría que encargarse, para ser práctico, de mucho más. Debería, por ejemplo, poderse imprimir la lista completa de números generados a partir de un valor inicial dado, o presentar la lista de valores máximos y longitudes de caminos asociados con todos los enteros comprendidos en un cierto intervalo. Podría prepararse otro programa destinado a buscar los enteros que dan lugar a trayectorias progresivamente más largas, o valores de pico cada vez más altos. Hay muchas otras posibilidades.

Vale la pena explorar también variantes de la fórmula $3N + 1$ que utilicen constantes y coeficientes distintos. R. William Gosper y Richard Schoeppel investigaron el problema "$3N - 1$" cuando formaban parte del HAKMEM, y mostraron que era equivalente al problema $3N + 1$ con valores negativos de N. Todos los números que estudiaron desembocaban en uno de tres posibles bucles; el de mayor extensión comenzaba en $N = 17$ y tenía un período de 18 pasos.

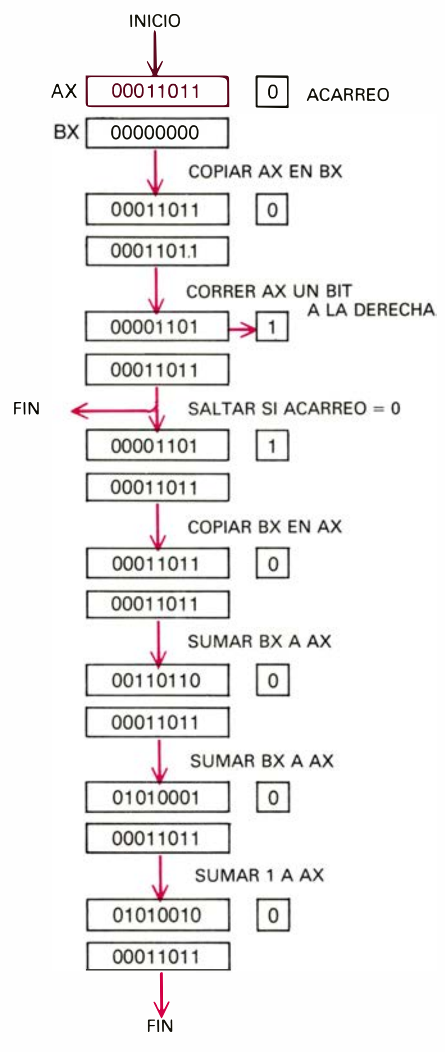
Un programa cuya finalidad única

N	LONGITUD DE LA TRAYECTORIA	VALOR MAXIMO
1	0	1
2	1	2
3	7	16
6	8	16
7	16	52
9	19	52
18	20	52
25	23	88
27	111	9232
54	112	9232
73	115	9232
97	118	9232
129	121	9232
171	124	9232
231	127	9232
313	130	9232
327	143	9232
649	144	9232
703	170	250.504
871	178	190.996
1161	181	190.996
2223	182	250.504
2463	208	250.504
2919	216	250.504
3711	237	481.624
6171	261	975.400
10.971	267	975.400
13.255	275	497.176
17.647	278	11.003.416
23.529	281	11.003.416
26.623	307	106.358.020
34.239	310	18.976.192
35.655	323	41.163.712
52.527	339	106.358.020
77.031	350	21.933.016

3. La serie de trayectorias máximas, hasta **N** *= 100.000*

N	LONGITUD DE LA TRAYECTORIA	VALOR MAXIMO
1	0	1
2	1	2
3	7	16
7	16	52
15	17	160
27	111	9232
255	47	13.120
447	97	39.364
639	131	41.524
703	170	250.504
1819	161	1.276.936
4255	201	6.810.136
4591	170	8.153.620
9663	184	27.114.424
20.895	255	50.143.264
26.623	307	106.358.020
31.911	160	121.012.864
60.975	334	593.279.152
77.671	231	1.570.824.736

4. La serie de valores de pico, hasta **N** *= 100.000*



5. Algoritmo para pedriscos

sea buscar números que no terminen cayendo en el bucle 4-2-1 puede ser muy "aerodinámico". Si se quieren ir examinando sucesivamente todos los números, a partir de 1, basta estudiar los impares. Todo número par es inmediatamente reducido a su mitad, y por tanto, la trayectoria que pueda generar habrá sido estudiada ya. Por idénticas razones, no hay necesidad de recorrer el camino completo hasta 1: tan pronto el valor de N cae por debajo del valor inicial, el candidato examinado puede ser eliminado. Reglas más eficaces todavía para reducir el campo de búsqueda fueron puestas a punto por William H. Henneman, quien de estudiante perteneció al grupo HAKMEM, del MIT, y hoy se encuentra en la Universidad de Boston.

Aunque carecemos por ahora de una demostración, tal vez haya indicios de explicación en un razonamiento heurístico más refinado que las tres ingenuas hipótesis mencionadas más arriba. Se hizo notar ya que en cualquiera de las fases del cálculo, N tenía iguales probabilidades de ser multiplicado por 3 como de ser dividido entre 2, y que ello inducía a pensar que su valor debería tender a quedar multiplicado por 3/2 en cada iteración. Lagarias señala, no obstante, que una cuarta parte de los enteros no sólo son divisibles por 2, sino también por 4; que un octavo de ellos son divisibles por 8, y una dieciseisava parte, por 16, etc. Tomando en cuenta la probabilidad de división entre todas las potencias de 2, la predicción resultante es que N debería decrecer según un factor 3/4 en cada iteración, predicción sustentada, según parece, por las comprobaciones empíricas.

Aun cuando se demuestre que la totalidad de enteros positivos acaban sucumbiendo en el bucle 4-2-1, los números granizo serán fuente de abundantes curiosidades. Tal vez las más intrigantes propiedades de estos números sean ciertas conspicuas pautas en la distribución de longitudes de trayectorias y de valores máximos alcanzados en ellas. Si números tan pequeños como 27 pueden mantener la pelota en juego durante 111 pasos, alcanzando alturas de 9232, se podría esperar que la longitud de trayectoria y el valor de pico crecieran rápidamente al crecer N. La verdad es que la longitud del recorrido crece muy lentamente; el aumento de valor máximo es más rápido, pero bastante errático de todas formas.

Entre los primeros 100 enteros, el camino más largo consta de 118 pasos (cuando $N = 97$); entre los primeros 100.000, la máxima longitud es de solamente 350 pasos (para $N = 77.031$). Así pues, al multiplicar N por un factor de 1000, la longitud de trayectoria solamente se triplica; la relación es, aparentemente, logarítmica. El registro máximo de 9232, alcanzado para $N = 27$, no se supera hasta $N = 255$, obteniéndose entonces un valor de pico de 13.120. Se registran nuevos máximos a intervalos bastante irregulares. La secuencia correspondiente a $N = 77.671$ alcanza la extraordinaria altura de 1.570.824.736.

Es fácil comprender que el valor de pico obtenido por una sucesión de pedrisco ha de ser, invariablemente, número par. Puede demostrarse también que solamente los valores impares de N pueden establecer nuevos récords de altura máxima (con la posible excepción de $N = 2$). No sé de razonamientos teóricos que exijan a N ser par o impar para conseguir registros máximos de longitud de trayectoria. No obstante, entre los 100.000 primeros enteros, los récords de longitud de trayectoria corresponden casi siempre a los establecidos por valores impares de N.

Los listados de longitud de trayectoria y valor máximo alcanzado en ella en un determinado intervalo de números presentan una desesperante amalgama de regularidad y desorden. Decididamente, no son aleatorios, pero sus pautas se resisten a interpretación. Ciertos valores máximos, por ejemplo, son mucho más frecuentes que otros, demasiado frecuentes incluso para explicarlos mediante ningún proceso estadístico. El ejemplo más destacable es 9232, número máximo alcanzado por vez primera cuando $N = 27$. De los primeros 1000 enteros, más de 350 tienen su valor máximo en 9232.

Igualmente peculiar es la distribución de longitudes de trayectorias. Pueden producirse trayectorias de longitud arbitraria (tomando las sucesivas potencias exactas de 2) pero, como antes sucedía, ciertos números se presentan más frecuentemente que otros. Además, tanto las longitudes de las trayectorias como los valores máximos parecen mostrar una fuerte tendencia a apiñarse. En 1976, Fred Gruenberger, de la Universidad estatal de California en Northridge, publicó una relación de tales apiñamientos; el máximo era una cadena de 52 números que tienen todos trayectorias de la misma longitud. ¿Podrán tener dos valores consecutivos de N la misma longitud de trayectoria y el mismo máximo? La cuestión puede decidirse por medios algebraicos, pero los lectores que prefieran la computación tal vez gusten de examinar las sucesiones correspondientes a los valores $N = 386$ hasta $N = 391$.

Un enfoque que pudiera arrojar algo de luz sobre el problema consiste en volverlo del revés. Supongamos verdadero para todos los números enteros positivos que al final sus sucesiones acaban en el bucle 4-2-1. Habrán entonces de formar una cadena sin ruptura por medio de la cual todo número de la sucesión infinita de números naturales queda conectado con el fondo del bucle. De acuerdo con esto, debería poderse invertir la función generatriz: comenzar con 1 y aplicar la transformación "en marcha atrás", a fin de ir engendrando cada uno de los números mayores. De existir algún número no alcanzable de este modo, remontando el río, por así decirlo, tal número no

podría desembocar en 1 como resultado final.

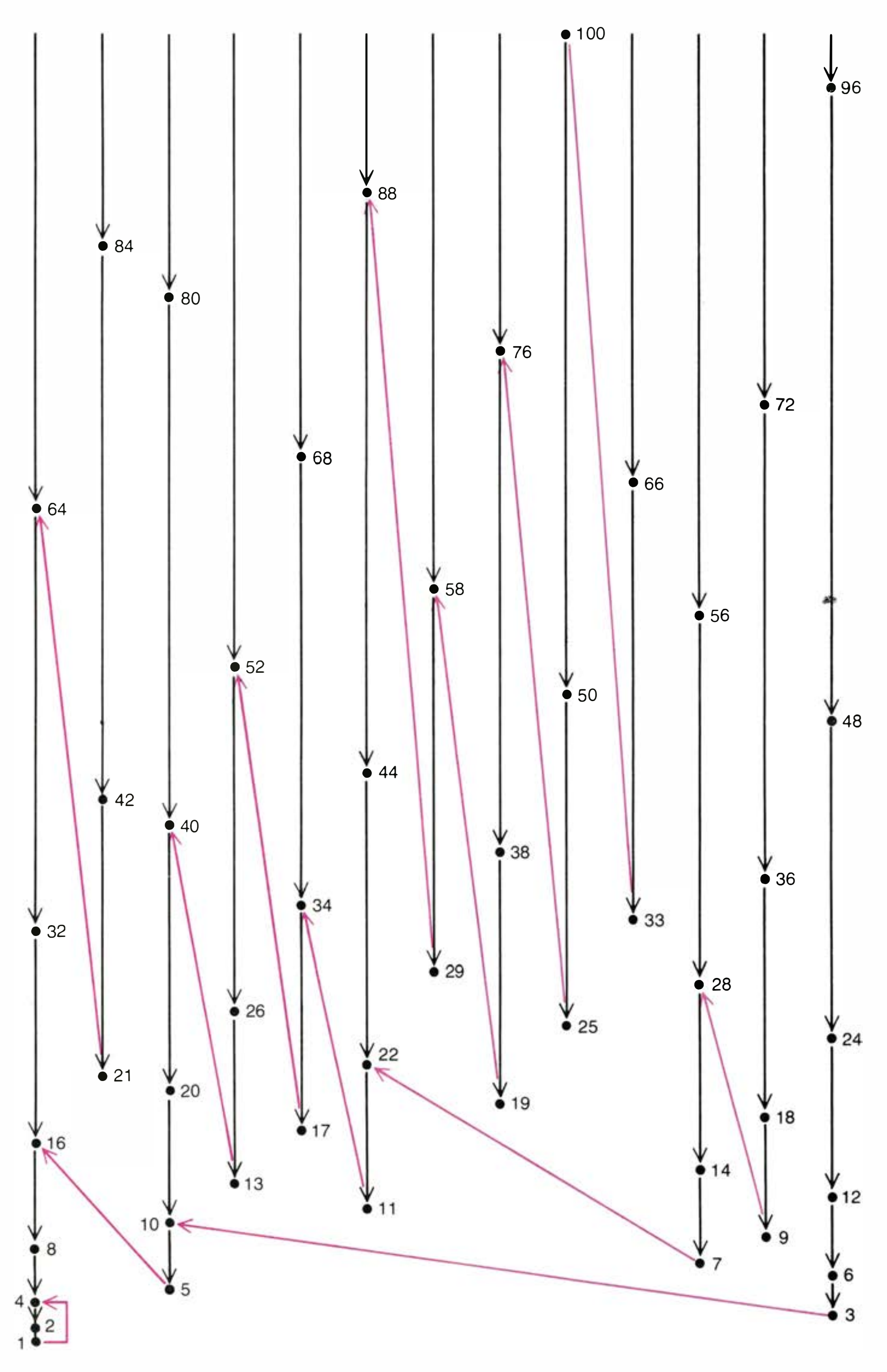
El método bien podría producir una solución general al problema de los números granizo, si pudiera llevarse a cabo hasta el fin. En realidad, el procedimiento indicado no es tan directo como parece. La función generatriz, en su forma normal, o directa, es determinística: cada valor de N, en cada punto del cálculo, tan sólo puede tener un sucesor inmediato. Si, por ejemplo, N es 40, el siguiente número será forzosamente 20. Pero al recorrer el camino marcha atrás hay ambigüedades. Si nos tropezamos con el valor $N = 20$ nos damos cuenta que solamente pudo haber sido generado por 40; en cambio, 40 tanto puede haber sido generado por 13 como por 80. La corriente se escinde al remontarla, y es preciso explorar ambas tributarias. Se produce una bifurcación en cada número de la forma $6K + 4$, donde K puede ser 0, o un entero positivo cualquiera.

Un sistema ramificado, como éste, tan sólo podrá explorarse hasta profundidad finita. Resulta preciso explorar cada rama individual hasta una profundidad determinada; alcanzado el límite, hay que desviar la atención a otra rama. Situando el límite en 100 se exploran 13 ramas, y queda confirmada la conexión de 49 números al sistema de arroyos y riachuelos numéricos. Con el límite en 1000 hay 84 ramas, pero solamente se cuentan 340 números. Una profundidad límite de 10.000 da pie a 1065 ramas, que enfilan y ensartan 4235 números. Observemos que más de la mitad de los números parecen situados en los intersticios que separan los ramales de la corriente. Al ir incrementando el valor límite, más y más números quedan atrapados, pero también más y más van quedando fuera. Si pudieran explorarse con profundidad infinita todas las ramificaciones del sistema, ¿acabarían todos los enteros por encontrar su lugar en él? No es otra la gran pregunta pendiente de solución.

En diciembre mencioné dos problemas de naturaleza combinatoria, cuya solución por métodos "no algorítmicos", con ayuda de estadillos electrónicos, no me parecía muy probable. Cierto número de lectores demostró que sí eran resolubles de este modo.

Recibí soluciones para la Torre de Hanoi de David Behar, John B. Jones, Jr., George Arthur Miller, J. B. Sladen, Alun Wyn-jones y otros. Las técnicas que aplicaron fueron similares. Lo que se hizo fue desensamblar un algoritmo donde los discos de numeración impar circulan en sentido de las agujas del reloj, mientras los de numeración par lo hacen en el contrario. De esta forma, una secuencia temporal de instrucciones queda convertida en una formación espacial.

Behar, Miller y Sladen resolvieron también el problema de las ocho reinas. La dificultad principal reside aquí en desandar lo andado cuando falla una solución en fase de desarrollo. Yo había supuesto que sería preciso llevar alguna forma de registro de los intentos fallidos, pero no es así. D. H. Fremlin, en un comentario al planteamiento y resolución de problemas mediante estallidos, señaló "un método que generará por turno todos los movimientos, sin otra memoria que la ya desplegada... Las fórmulas son complicadas, y de ejecutarse en forma natural conllevan algoritmos secuenciales, pero todas estas subrutinas de 'búsqueda' pueden quedar plasmadas en el estallido mediante cálculos a realizar *in situ*".



6. Arbol generado por inversión de la función generatriz de pedriscos

Taller y laboratorio

Las farolas del alumbrado público iluminan ciertos conceptos fundamentales de la mecánica cuántica

Jearl Walker

En lo que respecta al estudio de los átomos y las moléculas, la espectroscopía ha constituido para los físicos algo así como la piedra de Rosetta. Al principio no fue más que un confuso embrollo de colores, luego se sumergió en un misterio de energías y longitudes de onda y, finalmente, emergió para abrirse paso, en tres etapas sucesivas, hasta la mecánica cuántica. Aplicable a casi todos los campos de la ciencia y la técnica, proporciona también a los aficionados una posibilidad de vislumbrar la estructura atómica. Para investigar en torno a las transiciones atómicas, origen de la zona visible del espectro, hay varios experimentos sencillos. Por ejemplo, los espectros del sodio y del mercurio pueden observarse con lámparas de laboratorio y también con dos tipos muy comunes de farolas públicas, las de vapor de sodio y las de vapor de mercurio.

El espectroscopio constituyó, durante mucho tiempo, un artefacto de laboratorio clásico, aunque costoso. Consta de tres componentes. Con la luz procedente de un foco luminoso, una lente y una rendija forman una onda plana. Luego, mediante un prisma o una red de difracción esa luz se descompone en los colores que la integran. Por último, un pequeño anteojo permite al observador contemplar la dispersión con detalle. El espectro puede fotografiarse montando una cámara en el anteojo por medio de un adaptador.

Se acaba de idear un procedimiento más sencillo y barato mediante el cual pueda acometer estudios de espectroscopía el mismo aficionado. Leonard H. Greenberg y Thomas Balez, de la Universidad de Saskatchewan, describen unos ensayos en los que emplean una red de difracción barata montada ante el objetivo de una cámara de revelado instantáneo. Esa red de difracción es una reproducción fotográfica y consta de unas 600 rayas por milímetro. Como fuentes luminosas emplean lámparas de laboratorio rellenas de mercurio, neón, helio y criptón.

Richard L. Bowman, de la Universidad estatal de Kent, sugiere dirigir ese sencillo espectroscopio hacia la luz de las farolas públicas, y así estudiar los espectros de emisión de los gases que contienen. Aconseja, asimismo, registrar los espectros en película en blanco y negro para que cada fotografía pueda montarse en un proyector corriente. Con esa proyección se amplían los detalles del espectro y un grupo de alumnos puede verificar sus características valiéndose de una regla graduada.

Richard Breslow, de la Universidad de Connecticut, ha señalado que el espectro de una farola de sodio ofrece una característica curiosa: presenta una franja oscura en la parte amarilla del espectro, precisamente donde más luminoso es el espectro de una lámpara de sodio de laboratorio. Esa franja oscura se distingue con una red de difracción barata y a simple vista.

No hace mucho, David A. Katz, del Community College de Filadelfia, me remitió sus proyectos para el estudio de diversas fuentes de luz con una cámara de 35 milímetros o con una cámara vídeo (que es muy útil para clases numerosas). Emplea dos tipos de fuentes: una lámpara de laboratorio clásica, provista de un tubo de descarga gaseosa, y un artefacto que él mismo construye con útiles de laboratorio. En esta última fuente, se insufla aire comprimido, a través de una solución de una sal metálica y agua, en un frasco inclinado. Este, que se llena de iones metálicos suspendidos en aire, tiene un apéndice al que se conecta una manguera que conduce los iones hasta la llama de un mechero Meker, donde se excitan por choque con otras moléculas en un medio caliente. Luego, al desexcitarse, emiten su espectro característico.

El dispositivo para rociar los iones dentro del frasco puede construirlo cualquiera que posea una mínima habilidad en el soplado del vidrio. Se compone de dos tubos de este material, uno interior, de seis a ocho milímetros de diámetro, y otro, exterior, de 12 a 14 milímetros. Por debajo de ese dispositivo de tubos concéntricos sale un tubo capilar de un milímetro de calibre. Se pretende que el aire comprimido (procedente de una botella o de la red del laboratorio) sople por entre los tubos; así, al estrecharse el recorrido de salida, se creará una depresión en el tubo capilar. Si éste se halla sumergido en una solución de una sal metálica, la aspiración arrastrará iones con la corriente de aire.

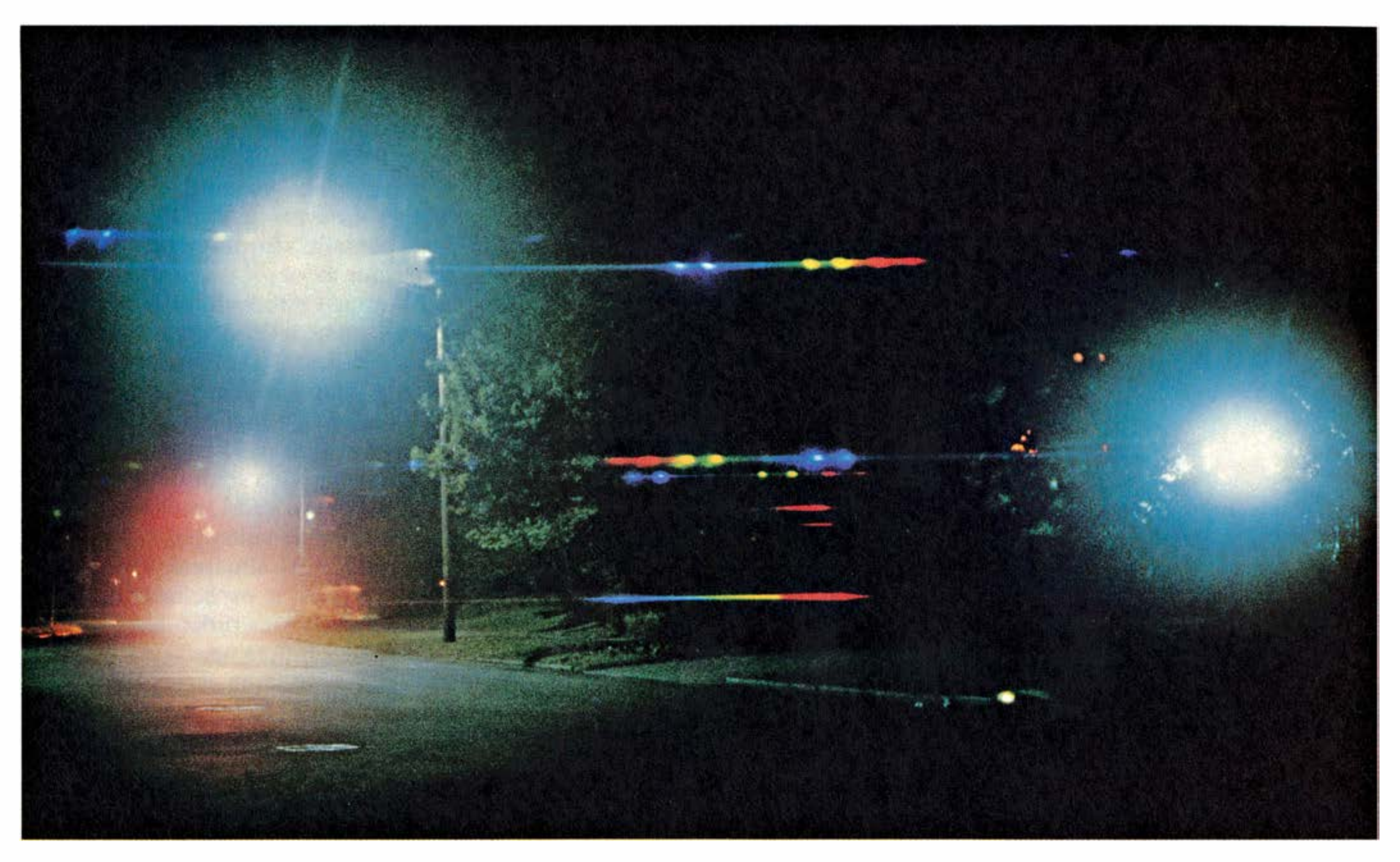
Katz recomienda construir primero el tubo interno. Para ello ensancha uniformemente un extremo y estrecha el otro hasta un diámetro pequeño. Luego, une el tubo capilar al costado del mayor de los tubos, donde introduce el tubo interno y une el extremo ensanchado de éste a un extremo del tubo mayor. La unión debe ser estanca al aire.

Se estrecha, a continuación, el otro extremo del tubo externo. Esta punta debe hacerse un poco mayor y más amplia que la del tubo interno. Con otro tubo de vidrio se forma una conexión flexible que se une al rociador. Este tubo poseerá un conectador de manguera adecuado, de modo que la manguera de caucho de la toma de aire ajuste sin huelgos. Se recuece el vidrio. Se abre el extremo inferior del tubo capilar y se amola a escuadra.

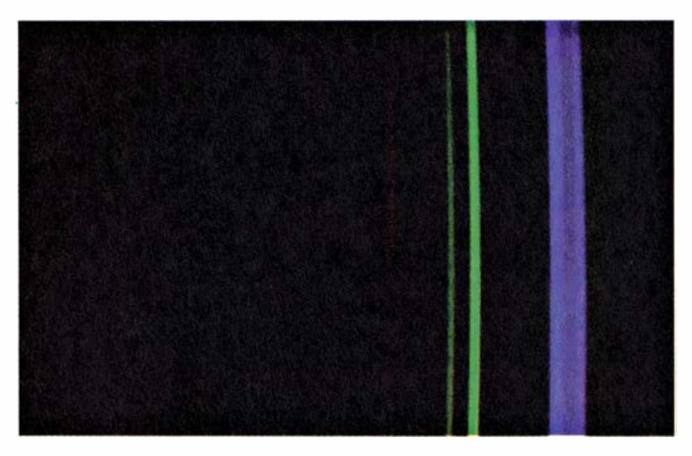
La capacidad del frasco debe ser de 500 a 1000 mililitros. A su apéndice se sujeta una manguera de caucho de unos 15 centímetros de largo y por uno de diámetro. Manguera que se fija con cinta adhesiva a uno de los respiraderos del mechero. Los demás respiraderos se cierran con cinta adhesiva para que la única entrada de aire al mechero proceda de la manguera. La alimentación del mechero con gas natural se hace de la forma acostumbrada. Así, el dispositivo suministra a la llama aire mezclado con un ion metálico.

Katz emplea redes de difracción por transmisión en forma de láminas de 22×28 centímetros. Sirven otras parecidas. Fija la lámina con pegamento a la visera de un objetivo fotográfico. Vale cualquier visera hecha de plástico, caucho o cartón. Cubre el borde exterior de pegamento y, luego, la coloca sobre la red. Cuando el pegamento se seca, recorta la parte de la red sobrante del borde de la visera.

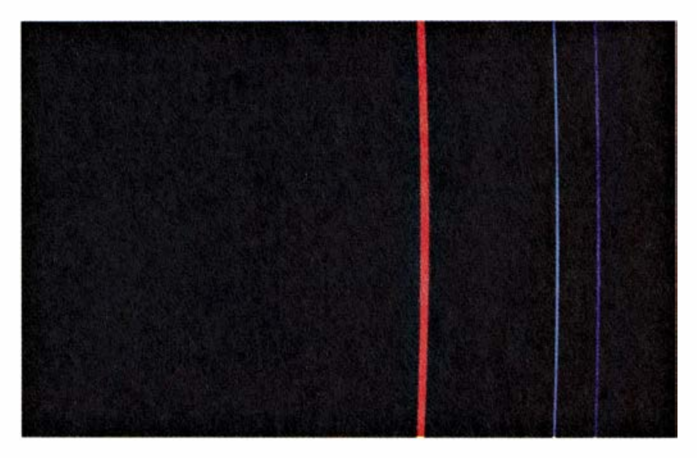
La cámara se monta en un trípode o en cualquier otro soporte estable. Se coloca a unos cuarenta o cincuenta centímetros de una fuente luminosa. Debe disponerse un fondo oscuro para eliminar los reflejos de los demás objetos de la habitación. Además, la cámara no se



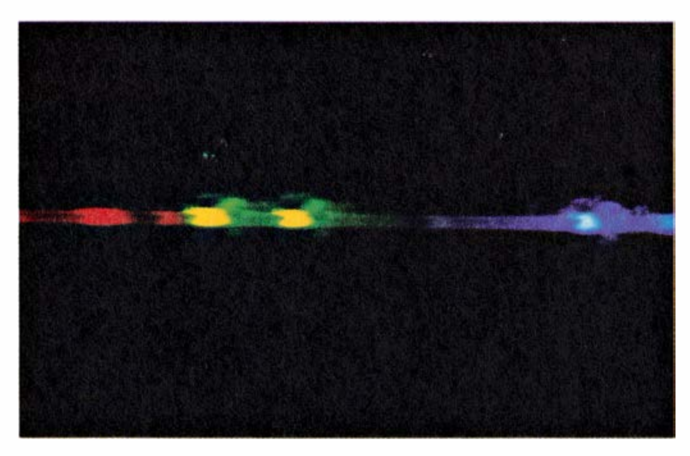
1. Vista de una calle con espectros



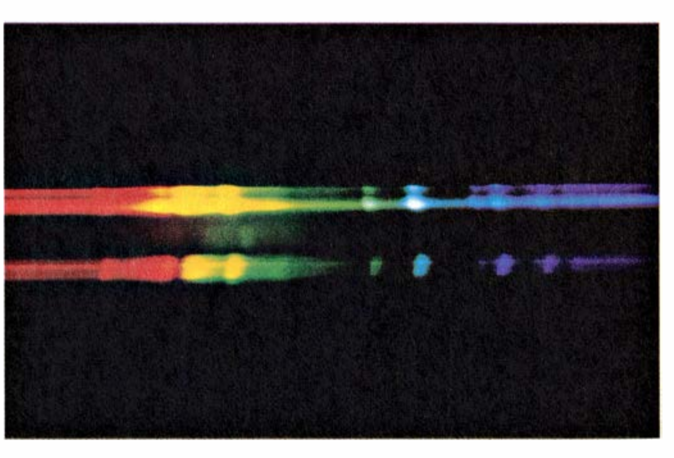
2. Espectro del mercurio tomado por David A. Katz



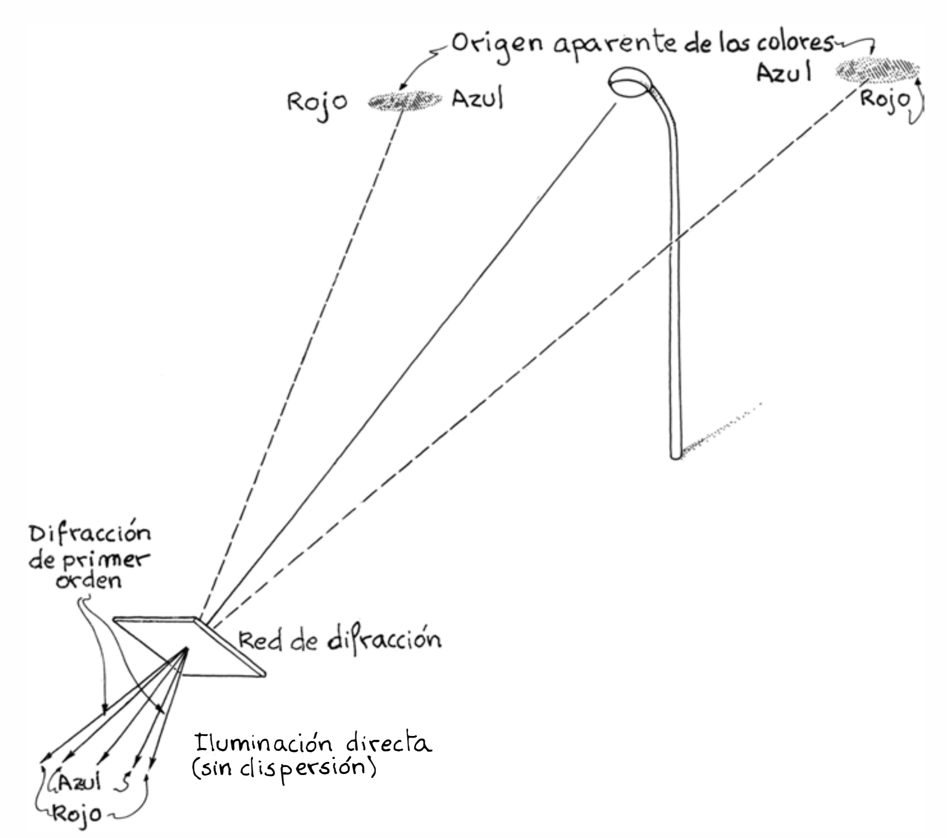
3. Espectro del hidrógeno tomado por Katz



4. Espectro de una farola de mercurio



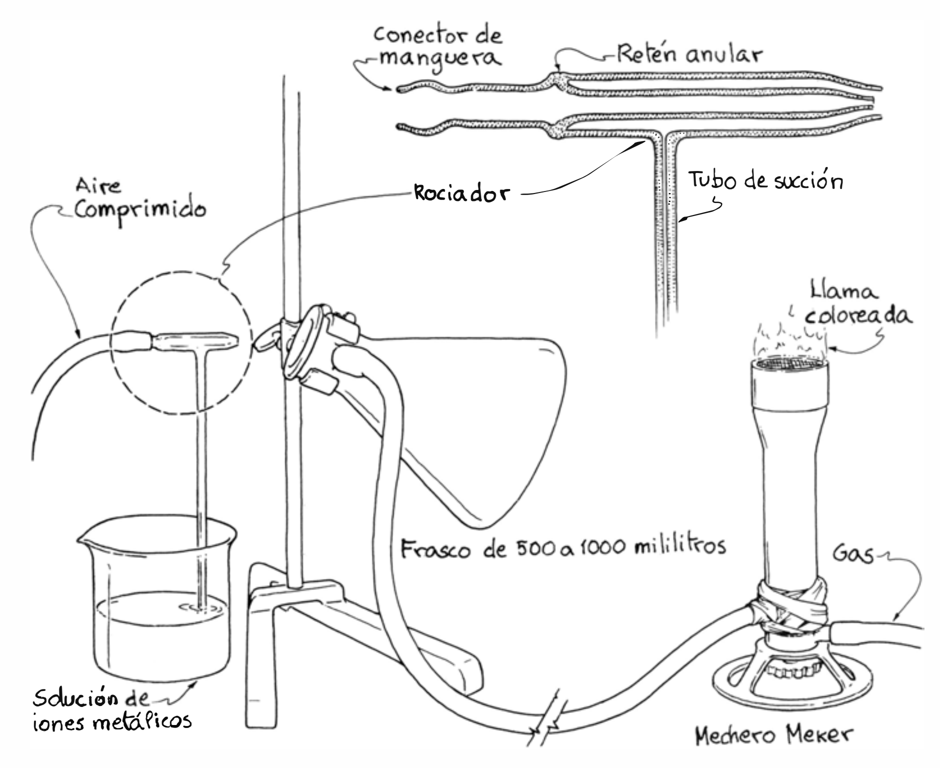
5. Espectros de varias farolas



6. Modo en que se dispersa el espectro producido por una farola

sitúa frente por frente al manantial, sino desalineada unos 30 grados. (De lo contrario, la red dejaría pasar la luz hacia la película sin dispersarla.) Así queda la película expuesta a la figura de difracción de primer orden que produce la red: la luz se distribuye de un lado a otro de la película en función de su longitud de onda.

En la cámara la apertura debe ser grande. El tiempo de exposición correcto sólo puede determinarlo la experiencia, aunque posiblemente se halle entre 1/4 segundo y 10 segundos, dependiendo de la luminosidad de la fuente y de la sensibilidad de la película. Katz recomienda Ektachrome o Ektachrome de gran sensibilidad. En lugar de un aparato fotográfico puede emplearse una cámara video, con lo que el espectro podrá transmitirse a una pantalla de televisión o a un proyector vídeo.



7. Dispositivo de Katz para espectroscopía de iones metálicos

He estado estudiando algunos espectros atómicos valiéndome de una red de difracción sencilla (530 rayas por milímetro), colocada delante de un objetivo montado en una cámara reflex de objetivo simple de 35 milímetros. Utilicé una película en color de sensibilidad ASA 400, pero quizá resulte interesante probar con la nueva película en color ASA 1000. Al situarme cerca de la fuente, las rayas espectrales de la difracción de primer orden tendían a solaparse: la primera aparecía considerablemente grande en mi campo visual. Retrocedí para estrechar las rayas, con lo que éstas se hicieron más mortecinas. Puse la cámara con la apertura máxima y tiré fotos con un gran número de tiempos de exposición, comprendidos entre 1/4 segundo y 25 segundos, en la esperanza de atinar con la exposición correcta. (Con una cámara de revelado instantáneo me hubiera ahorrado no poco de este trabajo de tanteo.)

Fotografié primero lámparas de laboratorio de hidrógeno, sodio y mercurio. Luego, una farola de vapor de mercurio. Esta emitía luz en todo el espectro visible, pero encima de ese fondo distinguía varias rayas luminosas. Aprecié azul oscuro (violeta), verde pálido, verde claro, amarillo claro, amarillo pálido (o rojo apagado) y rojo claro. (Los cinco fotogramas precedentes no recogen el detalle de los originales, ni de lo que se percibe cuando se contempla directamente la fuente a través del retículo.) La mayoría de esas emisiones provienen de la desexcitación de átomos de mercurio. La emisión roja se debe a un luminóforo de la lámpara que refuerza su potencia en el extremo rojo del espectro.

El espectro del sodio muestra un conjunto de colores diferente. En una lámpara de laboratorio, la emisión más intensa se encontraba en el amarillo, para una longitud de onda de unos 589 nanómetros. La emisión de una farola de sodio resultó distinta, en parte porque esas lámparas contienen una pequeña cantidad de xenón y mercurio, que intervienen en el espectro. Otra diferencia notable, señalada por Breslow, se refiere a la franja oscura del centro del amarillo.

Esta diferencia en la luz amarilla que

dan ambos tipos de lámpara procede de sus presiones gaseosas. En la lámpara de laboratorio la presión es de 0,02 atmósferas aproximadamente; en la farola, en torno a 1,5 atmósferas. En los dos casos los átomos de sodio se excitan por choque y emiten fotones de luz amarilla intensa con una longitud de onda de 589 nanómetros. Con una presión baja, cualquier fotón goza de buenas posibilidades de escapar de la lámpara sin que lo absorba otro átomo de sodio que encuentre en su camino. Si la presión es alta, el fotón tiene mayor probabilidad de ser absorbido por otro átomo y no escapar. Este fenómeno se conoce como auto-absorción.

Con todo, en la parte del espectro vecina a la franja oscura se ve luz amarilla a causa de las frecuentes y violentas colisiones entre los átomos. Una de estas puede alterar brevemente la energía de los electrones de un átomo. Si durante ese tiempo el átomo se desexcita, la luz presenta una longitud de onda un poco mayor o un poco menor que la luz amarilla normal. Esa luz no la puede absorber otro átomo de sodio, razón por la cual escapa de la lámpara sin auto-absorción.

Cuando el espectro del sodio de una lámpara de laboratorio se estudia en un espectrómetro de calidad, la emisión amarilla aparece en dos longitudes de onda: 589,0 y 589,6 nanómetros. Este doblete interesa porque es la prueba del sutil juego de los números cuánticos de un electrón del átomo de sodio.

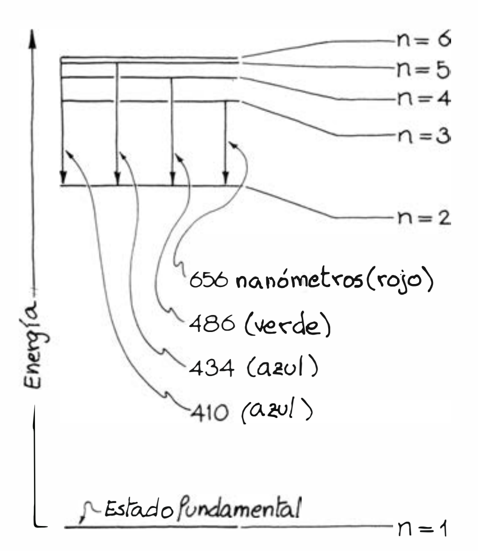
Reunir espectros sin saber nada de sus causas es como coleccionar mariposas sin conocer su biología. Durante dos siglos, los origenes de los espectros atómicos ocuparon las mentes de algunos de los físicos más eminentes. Su interpretación fue un proceso lento, que a menudo burló al mismo sentido común. Hacia la mitad del siglo XIX los estudiosos habían advertido una particularidad en la luz emitida por la combustión de elementos puros. Así, si bien la luz que emite el carbón que se quema en una chimenea muestra todo el espectro visible, un elemento puro produce únicamente determinados colores. Por ejemplo, la combustión del sodio da una luz amarilla peculiar. Todos los elementos que se investigaron presentaron su conjunto de rayas espectrales propias, fenómeno éste que desconcertaba a los precursores en estas investigaciones.

En 1885 Johann Jakob Balmer tropezó con una fórmula concisa que se ajustaba a las rayas del espectro visible del hidrógeno. Esa fórmula relacionaba la frecuencia (o longitud de onda) de las rayas con determinados conjuntos de números enteros. Uno de éstos era el 2; el otro, un entero mayor que 2. Ni Balmer ni nadie tenían la más ligera idea de por qué servía aquella fórmula.

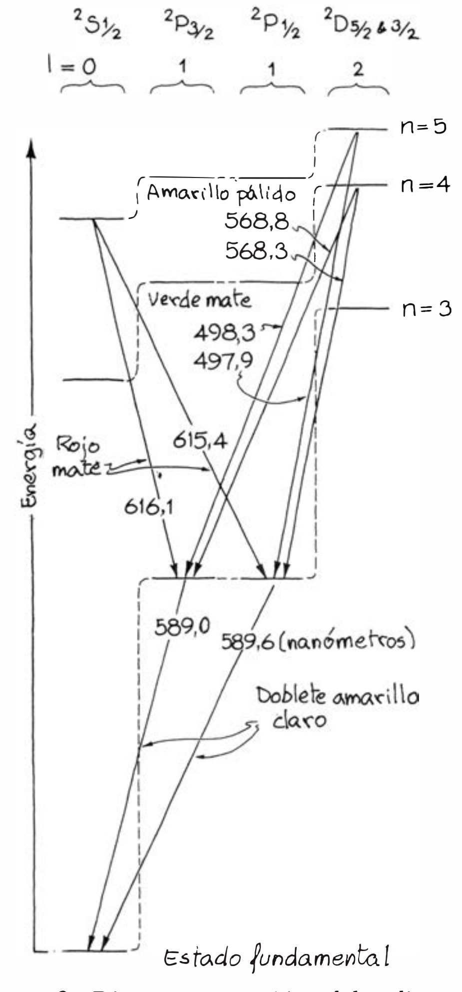
En 1911 Ernest Rutherford probó que un átomo constaba de un centro sumamente denso, y cargado positivamente, rodeado a cierta distancia de partículas cargadas negativamente. En la actualidad ya sabemos que ese centro, o sea, el núcleo del átomo, se compone de protones y neutrones y que las partículas circundantes son electrones. Cuando un electrón pierde energía y se pone a orbitar más cerca del núcleo, se emite luz; pero, ¿por qué se ven sólo ciertas rayas en el espectro? Según lo que se sabía de electromagnetismo en los tiempos de Rutherford, un electrón que perdiese energía no tardaría en caer sobre el núcleo. Si así fuera, todos los átomos colapsarían y no existiría el universo que conocemos.

El primer golpe de suerte en la resolución del misterio llegó en 1913, cuando Niels Bohr estudiaba el modo de aplicar al átomo de hidrógeno la nueva idea de "cuanto de acción". Con anterioridad, Max Planck había allanado el camino al introducir la cantidad que ahora llamamos constante de Planck, cuya finalidad era crear un modelo matemático para los emisores de luz atómicos situados en la superficie de un objeto caliente, carbón en combustión por ejemplo. La aplicación de ese modelo al átomo de Rutherford era una tarea ardua, aunque posible. Sin embargo, la explicación de los espectros atómicos se consideraba fuera de todo alcance.

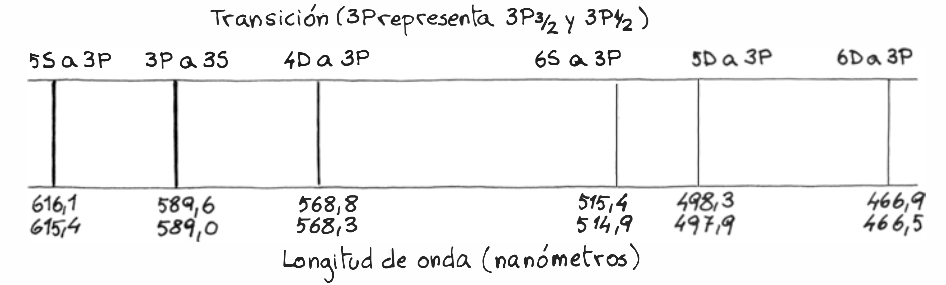
Cierto día un amigo sugirió a Bohr que examinara la fórmula de Balmer. Más tarde Bohr afirmaría: "En cuanto leí la fórmula de Balmer lo vi todo claro inmediatamente". En la fórmula aparecían números enteros, lo que insinuaba que alguna característica del electrón del átomo de hidrógeno estaría limitada a determinados valores función de esos números. Y, simplemente, el electrón no caía al núcleo porque ello violaría tal limitación. Bohr descubrió que, cuantificando el momento cinético de un electrón en órbita alrededor del núcleo, su energía quedaba entonces limitada a determinados valores establecidos por un entero al que él llamó número cuántico *n*. Si el electrón orbita próximo al núcleo, su energía tendrá un valor bajo, dentro del conjunto permitido. Una órbita mayor requiere un valor de *n* mayor. El electrón puede saltar de una energía permitida a otra, cambiando así el tamaño de su órbita, pero evidentemente no puede tener ninguna energía de valor intermedio, ni orbitar con un radio intermedio. Además, está prohibido el salto descendente al núcleo en sí. Como la energía que puede adquirir el electrón tiene un valor inferior definido, existe también un límite a la pequeñez de la órbita.



8. Diagrama energético del hidrógeno



9. Diagrama energético del sodio



10. Espectro del sodio

El electrón puede saltar de una energía a otra mayor únicamente absorbiendo energía del exterior del átomo. Tal energía puede aportarla un choque con otra partícula o átomo; también puede conseguirla absorbiendo luz. Tras excitarse, el electrón no tarda en desexcitarse liberando energía. Para librarse de esa energía, el átomo ha de chocar con algo, si no, emitirá luz. Cuando un átomo de hidrógeno pierde energía por emisión luminosa, dicha emisión forma el espectro luminoso que se ve en la luz que arroja una lámpara de hidrógeno.

La figura 8 presenta las energías permitidas del átomo de hidrógeno según el modelo de Bohr. El nivel más bajo, que recibe el nombre de estado fundamental, tiene un número cuántico n de 1. Los otros niveles se identifican con valores de n mayores. La fórmula de Balmer (en su forma original) abarca los niveles segundo y superiores. Por ejemplo, si un átomo de hidrógeno se excita al tercer nivel, enseguida se desexcita a cualquiera de los dos niveles más bajos. Supongamos que lo haga al segundo; en tal caso, para saltar, deberá deshacerse de la cantidad de energía exacta.

Con antelación al trabajo de Bohr, Albert Einstein había expresado la relación fundamental entre la energía y la frecuencia de la luz: la energía es igual al producto de la frecuencia por la constante de Planck. Por ejemplo, cuando el hidrógeno se desexcite desde el tercer nivel al segundo, deberá emitir luz de una cierta frecuencia al objeto de perder la cantidad de energía correcta. Como la frecuencia de la luz es inversamente proporcional a su longitud de onda, ésta se halla también determinada. El resultado de tal desexcitación se manifiesta en el espectro mediante una raya roja de una longitud de onda de 656 nanómetros.

En una lámpara los átomos están constantemente chocando y excitándose a todos los niveles que figuran en el diagrama. Las numerosas transiciones descendentes implican que diferentes cantidades de energía se liberan en forma de luz. O sea, cuando, por alguna razón, los átomos de hidrógeno se excitan sale una luz de diferentes longitudes de onda. Lo interesante es que cabe esperar sólo determinados valores. En la parte visible del espectro, además de la roja, hay tres emisiones, asociadas cada una a un salto descendente al segundo nivel. Del cuarto nivel sale una longitud de onda verde; del quinto, una azul; del sexto, otra azul de longitud de onda más corta. Estas emisiones caen las cuatro en el espectro visible. Otros saltos descendentes dan por resultado emisiones invisibles en las partes ultravioleta o infrarroja del espectro.

El modelo de Bohr para el átomo de hidrógeno supuso una revolución, por cuanto introducía en el átomo la cuantificación del movimiento y de la energía; sin embargo, era erróneo. De hecho, la correspondencia entre sus predicciones para el espectro del hidrógeno y el espectro observado fue casi accidental. Este modelo lo basó Bohr en la hipótesis de que el electrón describe una circunferencia en torno al núcleo, pero no hay motivos para que no puedan existir otros tipos de órbitas.

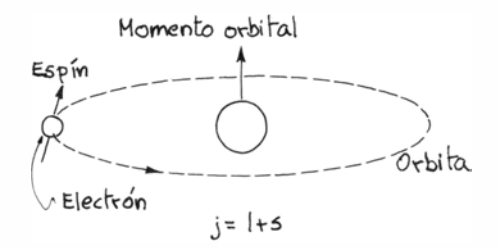
Investigaciones posteriores de Arnold Sommerfeld trataron de órbitas tridimensionales. Si bien el modelo de átomo de hidrógeno parecía, en aquella época, plenamente aceptable, fracasó cuando se quiso aplicar a otros tipos de átomos. Ni siquiera pudo explicar los sencillos espectros producidos por lámparas de laboratorio y farolas callejeras. Evidentemente algo fallaba.

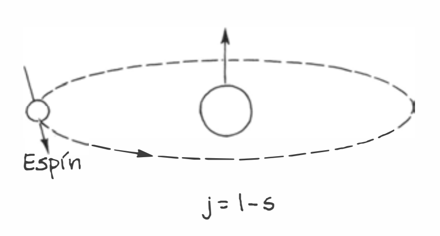
Un modelo mecano-cuántico válido para el átomo de hidrógeno surgió en 1926 cuando Erwin Schrödinger planteó la ecuación de onda que lleva su nombre. Esa ecuación venía a ser una repetición de la expresión que los físicos precuantomecanicistas habían estado empleando para las ondas de toda naturaleza. La ecuación de Schrödinger resultaba chocante porque no describía una onda en el agua, o en cualquier otro medio material, sino una onda en lo que recibe el nombre de campo escalar. La ecuación fue un acierto magnífico pues se acomodaba perfectamente a los espectros atómicos, pero el significado físico de campo escalar resultó tan escurridizo que muchos científicos consideraban aquello como una teoría antinatural.

De modo ordenado, de la ecuación fueron brotando nuevos números. Todo electrón de un átomo posee un número cuántico principal n (similar, en cierto modo, al n de la teoría de Bohr), un segundo l, que restringe el valor del momento cinético de la órbita del electrón alrededor del núcleo, y un tercero que restringe la orientación de ese momento cinético. Los dos primeros facilitan la interpretación de los espectros de las lámparas de laboratorio y de alumbrado público.

Una vez determinado el valor de n, el valor de l puede ser 0 o cualquier entero hasta $n-1$. Aquí es necesario recurrir a otro parámetro que limita el momento angular interno del electrón. Todo electrón se comporta cual si fuera una minúscula peonza que girase en torno a un eje. El momento cinético correspondiente a ese movimiento, que recibe el nombre de espín, lo da el número cuántico s.

Debo advertirles que todos estos números cuánticos, esquemas, diagramas energéticos, etc., no son más que modelos matemáticos asociados al átomo. No constituyen una realidad en el sentido corriente. Nadie observará jamás un electrón directamente, ni percibirá la cuantificación de sus órbitas y energías cuando esté en un átomo. Lo cierto es que un átomo no es un mero sistema solar en miniatura, ni el espín es meramente el momento cinético de una peonza diminuta. Según se ven hoy las cosas, el electrón es un punto infinitesimal al que pueden atribuírsele ciertas





11. Combinaciones de números cuánticos

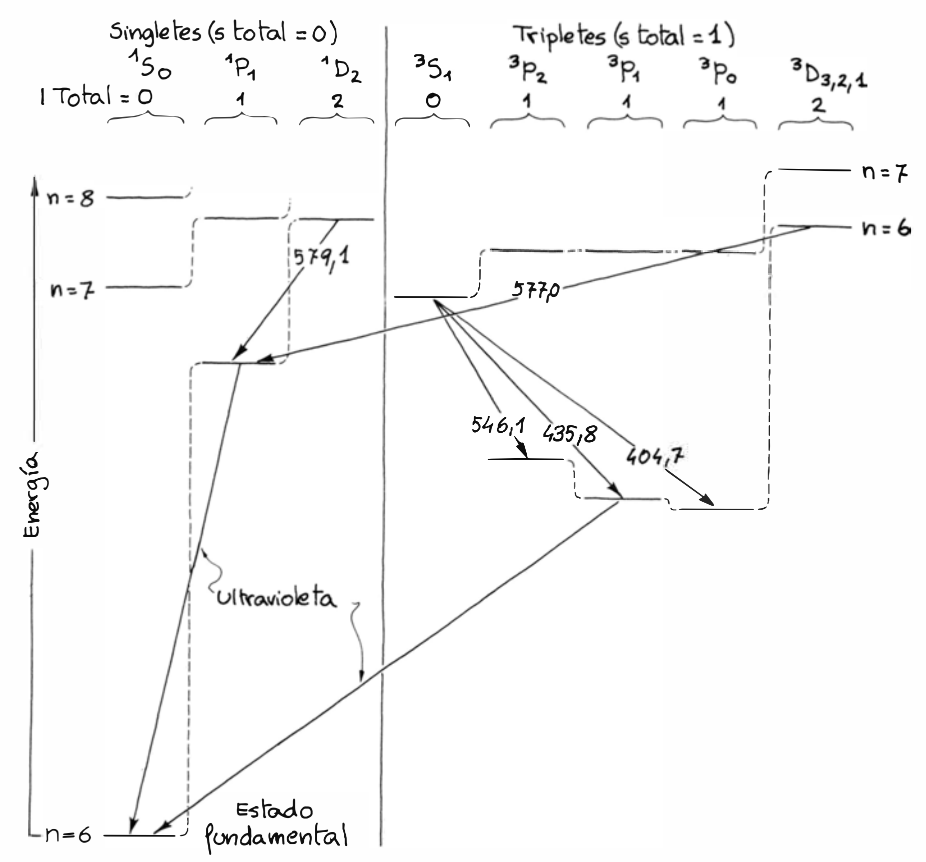
características, entre ellas carga y espín.

Podemos valernos del espectro del sodio para poner de manifiesto de qué modo aparecen los números cuánticos en los espectros atómicos. Un átomo de sodio neutro posee 11 electrones. Diez de éstos no desempeñan papel alguno en la interpretación del espectro porque forman lo que se llaman capas saturadas. El concepto de capa saturada se basa en la tesis mecánico-cuántica de que en un mismo átomo no puede haber dos electrones que tengan exactamente el mismo conjunto de números cuánticos.

El undécimo electrón se llama electrón de valencia porque se halla fuera de las capas saturadas. Cuando un átomo de sodio no está excitado, su electrón de valencia tiene un n de 3 y un l de 0. Para l hay dos valores posibles más (1 y 2); para que el electrón los adquiera debe recibir energía. En lugar de ello, podría recibir energía suficiente para saltar a un estado $n = 4$, o superior. Estas posibilidades, en lo que respecta a la energía, se representan en la figura 9.

El estado fundamental (en la parte inferior izquierda del diagrama) puede identificarse como $3^2S_{1/2}$. S es el símbolo que tradicionalmente representa al valor 0 del l del electrón de valencia. El número 3 es el valor de n. El supraíndice 2 proviene de una fórmula en la que interviene el espín. El subíndice introduce un nuevo número cuántico, que es una combinación de l y s. La importancia de este nuevo número radica en que limita el momento cinético total del electrón, o sea, la combinación de su momento cinético orbital y su momento cinético interno, o de espín. Los valores de j se calculan sumando y restando los de l y s. El mayor valor de j será la suma de los mayores valores de l y s; su menor valor será la diferencia entre l y s (ignorando los signos negativos que puedan salir). Los valores intermedios de j se encuentran restando 1 al mayor valor hasta llegar al menor valor. Para el estado fundamental el cálculo es sencillo. Para el mismo, l es 0 y s es 1/2. Por tanto, j sólo puede ser 1/2, que aparece como subíndice en el rótulo de columna.

Un átomo de sodio resulta más interesante si el átomo de valencia recibe alguna energía, de modo que (mientras prosigue en el estado $n = 3$) su l valga 1, dejando a j dos valores posibles. Si se suman l y s, j será 3/2; si se restan, 1/2. Físicamente estas dos situaciones se refieren a las orientaciones de las dos componentes del momento cinético del electrón. Si las componentes orbital y espinorial tienen el mismo sentido, j vale 3/2; j valdrá 1/2 si tiene sentidos contrarios.



12. *Diagrama energético del mercurio*

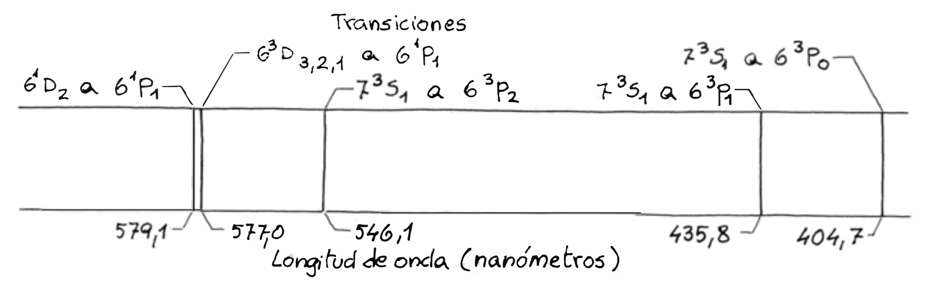
La única razón para introducir el número cuántico j es que la energía del electrón (y, por tanto, la del átomo) depende de él. Las dos orientaciones del momento cinético (las componentes orbital y espinorial) requieren energías diferentes para el electrón. Cuando los sentidos son contrarios la energía necesaria es menor; si iguales, un poco mayor. La manera de seguir de cerca la situación es considerar los valores de j.

En el diagrama energético los dos estados aparecen como los niveles más bajos en los rótulos $P_{3/2}$ y $P_{1/2}$. Como antes, el subíndice da el valor de j. P es el símbolo que tradicionalmente representa el valor 1 de l. La diferencia entre las energías de estos dos estados es tan leve que, en el diagrama energético, parecen estar al mismo nivel. Aun así, esa diferencia de energías responde de un doblete de rayas amarillas muy juntas que aparece en el espectro del sodio.

Supongamos que un átomo del sodio de una lámpara se excite al estado $P_{3/2}$. Rápidamente caerá al estado fundamental, con emisión de luz para librarse de la cantidad conveniente de energía. La longitud de onda de dicha luz es de 589,0 nanómetros, correspondiente a la parte amarilla del espectro. Si, por otro lado, el átomo se excitara al estado $P_{1/2}$, se desexcitaría emitiendo una cantidad de energía un poco menor, pues $P_{1/2}$ tiene un nivel energético ligeramente inferior. En esta ocasión, la longitud de onda de la luz será de 589,6 nanómetros, que corresponde a la segunda raya del doblete amarillo del sodio. La presencia de este doblete impone, pues, la invención del número cuántico j; si no, esas dos longitudes de onda amarillas levemente diferentes no se entenderían.

En el diagrama se representan varios de los demás estados de excitación que alcanzan los átomos de la lámpara. Tan pronto se excitan, los átomos regresan al estado fundamental por diversos caminos. Mientras la lámpara permanece encendida, los átomos vuelven a excitarse. De lo que resulta una emisión permanente de luz a las longitudes de onda originadas por todas las desexcitaciones posibles. He recogido algunas de las posibilidades del diagrama energético y el esquema del espectro del sodio.

Cuando el electrón de valencia se excita primero al nivel de $n = 5$ y $l = 0$, resulta una raya de emisión roja. Por motivos en los que no voy a entrar, el electrón no se limita a caer directamente al estado fundamental. En el diagra-



13. Espectro del mercurio

ma, el itinerario de desexcitación debe ser, a la vez, descendente y hacia otra columna de niveles. De aquí que el electrón pueda desexcitarse cayendo al nivel $P_{3/2}$, o bien al $P_{1/2}$. En un caso emite una longitud de onda de 616,1 nanómetros y, en el otro, una de 615,4. La diferencia entre ambas longitudes de onda débese a la diferencia entre las energías de ambos estados *P*. Y, así, los dos valores posibles de *j* en éstos acaban creando otro doblete de rayas espectrales muy juntas.

El análisis del espectro del mercurio es similar al del espectro del sodio. El átomo de mercurio neutro tiene cinco capas completas y dos electrones de valencia. Aunque la explicación de los números cuánticos para dos electrones es más trabajosa que para uno, los procedimientos son, más o menos, los mismos. En la figura 12 se representan los niveles energéticos del mercurio. Como antes, en la parte inferior izquierda se representa el estado de un átomo no excitado. En el diagrama aparecen más altos algunos de los estados excitados posibles. Los átomos de mercurio de una lámpara se ven, constantemente, excitados a esos niveles superiores; luego, se abren camino de regreso al estado fundamental mediante la emisión de luz.

Al analizar los niveles energéticos del mercurio, la diferencia principal estriba en que hay que centrarse en los *l* y *s* totales de los electrones de valencia. Estos totales se combinan para hallar el *j* total, el cual determina la energía. Para facilitar los cálculos el diagrama energético se divide en dos secciones que se rotulan singletes y tripletes. Bajo la denominación de tripletes se agrupan los niveles energéticos de los casos en que los dos electrones tienen los espines opuestos y, por ello, su *s* total es 0. Entonces, el *j* total de cualquiera de esos niveles no es más que el *l* total. Por ejemplo, el nivel más bajo rotulado *P* en esta sección corresponde al caso en que un electrón tiene un *l* de 0 y el otro lo tiene de 1. El *l* total será 1. Y, como el *s* total es 0, el *j* total valdrá también 1. En la sección de tripletes el lado izquierdo, rotulado *S*, tampoco presenta dificultades. Aquí, el *l* total es 0 y los espines de los electrones tienen el mismo sentido para dar una *s* total de 1. Por ello, el *j* total puede solamente ser 1.

Lo verdaderamente curioso se da en la otra parte de la sección de tripletes. Atenderé a los niveles más bajos, designados *P* en esa sección. Ahí el *l* total es 1, pero el *j* total admite tres valores posibles, asociado cada uno a una energía diferente. La suma de un *l* total y de un *s* total da un *j* total de 2. La sustracción da un *j* total de 0. Un valor intermedio del *j* total es 1. En el diagrama energético se muestran las energías asociadas a esos tres valores.

En el espectro del mercurio hay tres longitudes de onda, que dependen de esos tres niveles energéticos. Consideremos un átomo que se excite por choque hasta el nivel S_1 más bajo. Podrá desexcitarse cayendo a un nivel *P* de la misma sección del diagrama. (Al igual que en el diagrama del sodio, el recorrido de desexcitación debe ser descendente y hacia otra columna.) Como hay tres opciones de niveles *P*, el átomo podrá emitir luz de 546,1, 435,8 o 404,7 nanómetros. Dado el enorme número de átomos que se excitan dentro de la lámpara, siempre habrá algunos que estén cayendo del nivel S_1 siguiendo los tres itinerarios, produciendo una emisión ininterrumpida a las tres longitudes de onda. Así pues, la presencia en el espectro del mercurio de este triplete de longitudes de onda sólo puede explicarse con el número cuántico *j*.

Lo referido hasta aquí constituye una introducción al estudio de los espectros de las lámparas ordinarias. Vale la pena descubrir otras características de las lámparas de sodio y mercurio. Por ejemplo: analizar el espectro de una lámpara fluorescente y revestida de un luminóforo en su superficie interior. Hay lámparas cuya luz la emiten moléculas, y no átomos.

Libros

El Argar, manual de física, problemas de álgebra, ciencia y cultura y álbum de las corrientes fluidas

D. Fletcher, L. Bru, J. A. Navarro, J. Nubiola y M. G. Velarde

LA "CULTURA" DE EL ARGAR, por Vicente Lull. Akal Editor; Madrid, 1983. El libro, que corresponde a una tesis doctoral, está orientado hacia el estudio económico-social de la cultura argárica más que hacia el estrictamente arqueológico, pues es su propósito, como nos indica en la introducción, que sirva de modelo de estudio para las sociedades que únicamente conocemos a través de la arqueología, exponiéndonos sus principios metodológicos con larga disquisición sobre las escuelas funcionalista, estructuralista y marxista, pero como ha de basarse, por fuerza, en el análisis de los restos arqueológicos proporcionados por las excavaciones, comienza por esta tarea previa, lamentando las deficiencias de las excavaciones de fines del siglo pasado y primer cuarto del presente, que dificultan la utilización de su metodología.

Define el hábitat como el espacio donde se establecen las correlaciones económico-sociales y políticas entre la geografía y el grupo, y fija el espacio argárico desde los 36º 39′ 2″ a los 38º 18′ 18″ de latitud y desde 0º 45′ a los 4º 15′ de longitud oeste, quedando fuera los yacimientos, tradicionalmente considerados como argáricos, de Cabezo Redondo, Las Peñicas, Terlinques y El Campello, que entiende, coincidiendo con Aparicio Pérez, más próximos al Bronce Valenciano. La superficie total puede calcularse en unos 49.000 kilómetros cuadrados, con muy variada orografía y factores ecológicos concretos, que va señalando a través de la palinología, edafología y fauna, destacando que la circulación hídrica era más abundante que en la actualidad y, por tanto, mayor la extensión de bosque y la riqueza agrícola y faunística, siendo el factor humano fundamental en el proceso de deforestación y desertización del territorio.

La discusión tipológico-cronológica de los materiales se hace en el capítulo dos, empezando por la cerámica, en la que distingue los conjuntos procedentes de viviendas de los de las necrópolis; mediante múltiples cálculos, determina formas y tipos, acompañando abundantes gráficos y una tabla resumen de las 8 formas establecidas por Siret. Lamenta la falta de análisis de arcillas y pastas cerámicas, pues ellos podrían informarnos sobre la existencia o no de alfares que comercializaran sus productos o si éstos eran fabricados in situ. Sigue el estudio de los objetos de metal (cuchillos, puñales, espadas, alabardas, hachas), fijando la tipología de todos ellos y discrepando de las conclusiones de Blance, Ruiz Gálvez y Schubart, y advierte que las cuentas de collar, los botones en *V*, los brazaletes de arquero, los punzones, cinceles y puntas de flecha no pueden tomarse como fósiles directores, dada la amplia cronología que tienen. El resultado pormenorizado de los materiales le lleva a discutir la división en Argárico A y B, puesto que ello no se corresponde con la realidad, según nos prueba el autor.

Estudia los asentamientos, necrópolis y hallazgos aislados, agrupándolos por proximidad económica y no geográfica, siendo más importante la información que proporcionan las necrópolis que la de los poblados. La minuciosidad y abundancia de datos que se aportan hacen imposible entrar en detalles, pero sí podemos decir que este capítulo es un valioso instrumento de trabajo.

Establece luego sus "Conclusiones particulares" sobre los rasgos definitorios de la cultura argárica, sin considerarlas definitivas, ya que los resultados que obtiene deben entenderse como una opinión personal. Según él, los grupos ubicados en las depresiones son de agricultura intensiva, mientras que los grupos montanos y de los altiplanos lo son de agricultura extensiva secundada por la ganadería, la cual es más importante en la zona de Granada que en el sudeste, mientras que la minería es de primer orden entre los grupos almerienses-murcianos, y lo mismo en los grupos de la depresión del Guadalquivir. La base de la alimentación cárnica está en los ovicápridos y cerdos, aunque también se aprovecha el buey, ciervo, oso, jabalí, etc. El caballo puede considerarse como un bien de prestigio y relacionarse con el transporte y las comunicaciones. En cuanto a la caza, practicada en todo tiempo, se incrementa al final del período, tal vez a causa del bajo rendimiento agrícola.

Existen poblados básicamente mineros, mientras otros lo son agrícolas y agrícola-ganaderos, siendo pocos los que alternan estas actividades con la metalurgia local, lo que induce a suponer que los objetos metálicos hallados en algunos asentamientos de esta índole deben considerarse como llegados en régimen de intercambio.

El desarrollo de la metalurgia implica un abandono de la agricultura y una mayor deforestación en busca de fuentes de energía calórica, razones por las cuales, al agotarse los filones, la tierra no está en condiciones de mantener una población que había incrementado su demografía. La degradación debe situarse hacia el 1400/1300 a. C.; ante la imposibilidad de salvar la crisis, se abandonan la mayoría de los asentamientos, buscando algunos grupos la supervivencia en otras tierras.

Al referirse a los orígenes de la cultura argárica, destaca que todos los artefactos de ésta se encuentran en el horizonte eneolítico local. Coincide, pues, con las opiniones de Bosch Gimpera, Carriazo, Schüle y Aparicio, y en parte con Martínez Santa-Olalla, quien ve una evolución in situ con relaciones con grupos anatólicos, postura que parecen aceptar Blance y Schubart, para quienes es hacia el 1400 cuando se inician las influencias exóticas. Frente a ellos está la posición de Almagro, Bosch, Maluquer y Arribas, quienes ven un origen anatólico para la cultura argárica. Por nuestra parte, creemos que, junto a la evolución in situ de los elementos eneolíticos, habría que pensar en otros llegados de fuera, tales como los enterramientos en pithoi, modalidad de inhumación que el autor pretende explicar suponiendo una especialización alfarera en la fabricación de grandes urnas cuyo uso se generalizó para los enterramientos, pero no podemos olvidar que este sistema se extiende por toda la cuenca del Mediterráneo y que en el área granadina predominan las inhumaciones en cistas y fosas, posiblemente por su mayor alejamiento de los centros receptivos de las costumbres mediterráneas.

Desde el punto de vista cronológico, establece una fase de formación de la cultura, que sitúa entre 1900/1800 a. C.; el apogeo, entre 1700/1500, con momento de gran desarrollo y expansión entre 1650/1550 a. C., llegándose al final hacia el 1300, aunque admite que pudiera alcanzar el 1000 a. C. A la vista de estas fechas finales para la cultura de El Argar, opinamos que es necesaria una amplia investigación que nos muestre cuáles eran las formas de vida en el área argárica desde el 1300 hasta los aportes fenicios del s. VIII.

El propósito de la obra tiene sus precedentes. Hoy en día, raro es el estudio en el que no sólo se hable de formas y tamaños, sino también de procesos de elaboración, análisis, flora, fauna, áreas de dispersión de los productos, sacando de todo ello consecuencias de tipo socio-económicas y espirituales, faceta ésta que el autor deja de lado, omitiendo toda referencia artística, creencias, lengua, etcétera, preocupado tan sólo en el aspecto material, que no lo es todo en la vida de los pueblos.

El hecho de que los tres primeros capítulos ocupen 396 páginas de apretado texto, frente a las 42 del IV y último capítulo, nos viene a probar que en el campo de la prehistoria toda conclusión ha de basarse en el previo análisis de los restos materiales, por lo que no debe relegarse a un segundo término la labor de los técnicos excavadores; luego vendrán, como dice el autor, "las interpretaciones, hipótesis, sugerencias, pensamientos, pues sólo así nos acercaremos más a los sistemas de producción y reproducción de las comunidades prehistóricas", opinión que está en línea con la que mantuviera, hace más de cuarenta años, Martínez Santa-Olalla, al enfrentarse a los "pucherólogos" e insistir que el estudio de los materiales ha de tener como finalidad alcanzar los datos ergológicos necesarios para configurar la personalidad de una cultura. Nosotros mismos hemos esbozado la economía del pueblo ibérico, aprovechando cuantos datos disponemos de su cultura.

Finalizamos con una observación: Es de lamentar la deficiente presentación del libro, con un texto apretujado y mal impreso, que hace fatigosa la lectura a pesar del interés con que ésta se sigue, tanto por la abundante aportación de información como por los conceptos que expone el autor. (D. F.)

INTRODUCTION TO MODERN PHYSICS, por Elmer E. Anderson. Saunders Golden Sunbrust Series, 1983. E. Anderson es un conocido profesor americano que goza de un elevado prestigio por su labor docente. El libro que llega a nuestras manos es un buen ejemplo de ello, al tratar los conceptos fundamentales de la llamada *física moderna*, sin utilizar la matemática del mismo apellido, sin apartarse nunca del fenómeno físico y utilizando un lenguaje sobrio y ortodoxo. Estoy convencido que será de gran utilidad para los estudiantes del primer ciclo de nuestras facultades, tanto de físicas como de químicas e incluso de exactas, así como también para los que cursan ingeniería en sus diferentes niveles y especialidades.

Cada capítulo contiene una serie de problemas bien elegidos, cuyas soluciones aparecen en las páginas finales, y de ejemplos, que son en realidad problemas resueltos, aplicaciones a veces casi inmediatas de fórmulas, leyes y ecuaciones que han sido deducidas en el capítulo correspondiente. En general, utiliza el sistema internacional de unidades SI (mks) que facilita mucho la labor y el esfuerzo de los lectores. También cada capítulo cuenta con una bibliografía somera de libros fáciles de encontrar en cualquier biblioteca, e incluye un resumen que ayuda a fijar las ideas principales. Asimismo, se insertan fotografías, algunas "históricas", fechadas, de destacados hombres de ciencia y una lista completa cronológica de los premios Nobel de física desde el número 1, otorgado en 1901 a Wilhem Konrad Roentgen, por el descubrimiento de los rayos X, hasta el de 1981 distribuido entre los americanos Nicolas Bloenberg y Arthur Schawlow, por su contribución a la espectroscopía láser, y al sueco Kai Siegbahn por su aportación al desarrollo de la espectroscopía electrónica de alta resolución. Todo ello contribuye a que se haya conseguido un libro amable en el que los estudiantes que gusten, además, del proceso histórico de la física, leerán con agrado.

Por costumbre, el título de *física moderna*, se refiere al rápido desarrollo que los conceptos de la física han experimentado en el siglo XX, considerándose como *física clásica* todo lo concebido hasta el año 1900, lo que no significa que esté pasada de moda ni anticuada. La mecánica y la termodinámica, por ejemplo, siguen aguantando impertérritas el empuje de las corrientes actuales pese a estar edificadas sobre unos pocos principios, sorprendentemente, pocos. La astronáutica, en contra de la creencia vulgar, utiliza la física clásica, la mecánica newtoniana, pero su espectacular avance es debido a la enorme cota alcanzada por la tecnología de la electrónica, en particular de las computadoras, de las comunicaciones y también de los combustibles, pero no hay que olvidar que las trayectorias que se imponen a los satélites y a los vehículos espaciales las fijan los principios inmutables de la mecánica clásica. Los jóvenes aman la física moderna y hasta desprecian un poco la clásica. Es un pecado que todos hemos cometido. Recuerdo a este respecto que, durante nuestra guerra, nos reuníamos con nuestro maestro, Julio Palacios, para discutir problemas de física. Al tratar de elegir un tema, Velayos y yo, por ejemplo, sugerimos enseguida la mecánica ondulatoria. Palacios, casi con timidez, nos propuso, con cierta desilusión para nosotros, tratar un tema tan clásico como la propagación del movimiento ondulatorio. He de confesar que pronto empezamos a encontrar dificultades para explicar satisfactoriamente las cuestiones que, por no haber meditado de forma debida, nos parecían triviales. Aquello supuso una gran enseñanza para mí, incrementando mi devoción por la física clásica. No se puede penetrar en la física moderna sin un conocimiento claro de la física clásica.

El desarrollo de la física es un proceso evolutivo. Cada teoría que surge trata de explicar los hechos nuevos, que la experimentación ha puesto de manifiesto, pero debe dar cuenta, como caso particular, de lo anterior. La física clásica cubre el mundo para el que nuestras percepciones están preparadas, el mundo que conocemos, tanto en tamaño como en velocidades. La física moderna ahonda en ese mundo en el que nuestros órganos sensoriales no son capaces de observarlo. Así ocurre con la mecánica relativista, que engloba la de Newton cuando las velocidades presentes en el proceso están muy lejos de alcanzar la de la luz en el vacío. El principio de causalidad, que domina en la física clásica, ha habido que revisarlo. La palabra probabilidad es la piedra angular de la física moderna.

Quizás el mayor éxito de Anderson es el orden de exposición de las materias tratadas. Comienza con un capítulo sobre *Atomos y teoría cinética*, de evidente sabor clásico. Continúa con otro dedicado al *Descubrimiento del electrón*, recalcando dos hechos importantes, el que las cargas eléctricas aparecen en unidades discretas (es decir, están cuantificadas) y el de la existencia de partículas subatómicas, lo que conduce a presentar la primera familia de

ellas constituida, además de por el electrón, el protón, el neutrón y el electrón positivo, introduciendo la propiedad del espín.

El capítulo 3 está dedicado a la *Relatividad especial*, enunciando los dos postulados de Einstein y sus consecuencias. Introduce el tiempo como una cuarta dimensión a añadir a las tres presentes en la geometría euclídea, insistiendo en el papel fundamental que desempeña el *intervalo* en el problema de la simultaneidad, llamando la atención sobre el hecho de que, según su característica, puede haber un sistema de referencia en el cual los sucesos son simultáneos (ocurren en el mismo instante), pero no existe un sistema en el que los sucesos ocurren en el mismo lugar y al revés. Sigue un cuarto capítulo en el que aparecen la *equivalencia masa-energía*, el sistema de referencia *C*, caracterizado porque el momento total de las partículas es nulo, la creación y aniquilación de partículas y el efecto Doppler.

En el capítulo 5 se aborda la *Teoría cuántica de la energía electromagnética*, deduciendo la probabilidad de las transiciones espontáneas e inducidas entre niveles de energía, lo que conduce inmediatamente al láser, describiendo con algún detalle los de rubí y helioneón. *El efecto fotoeléctrico y el efecto Compton*, bases experimentales que confirmaron las ideas de Planck, junto a la naturaleza dual del fotón y el principio de incertidumbre de Heisenberg, así como la aniquilación y la creación del par electrón-positrón, llenan el capítulo 6. *El modelo atómico de Bohr*, los espectros ópticos y de rayos X y la conexión entre la teoría cuántica de la radiación y de la materia están magníficamente tratados en el capítulo 7. *Las ondas que acompañan al movimiento de los corpúsculos*, previstas por de Broglie para aclarar algunos puntos oscuros de la teoría de Bohr, confirmadas por la difracción de las mismas por los cristales, aparecen en el capítulo 8. La definición de una función de ondas para representar de forma matemática la naturaleza ondulatoria de las partículas, el concepto de paquete de ondas para resolver el dualismo clásico onda-corpúsculo, la correspondencia entre la velocidad de la partícula y la de grupo del paquete, así como la falta de significado físico de la velocidad de fase del mismo se explica con claridad también en este mismo capítulo.

El noveno lo llena la ecuación de Schroedinger, justificando su necesidad el problema de los valores propios y los estados cuánticos de energía de los sistemas. El capítulo 10 está constituido por *Aplicaciones* de lo consignado en el anterior, pozos y barreras de potencial, efecto túnel y el oscilador armónico. El 11, lo forma la *Aplicación de la mecánica cuántica a la teoría atómica*. Es muy denso, pues opera con los números cuánticos, los niveles de energía y la probabilidad de densidad electrónica, el principio de exclusión de Pauli, los momentos magnéticos elementales, la cuantificación espacial, la interacción espín-órbita, el factor de Lande y el efecto Zeeman. Está tratado con claridad en la forma tradicional. Los ejemplos y problemas que se incluyen ayudan a la mejor comprensión de lo expuesto.

La introducción a la parte de la física en la que más se investiga en la actualidad, la *física del estado sólido*, se trata con relativa amplitud en el capítulo 12. La teoría de los electrones libres en los metales, bandas de energía, semiconductores, uniones diversas, transistores, etcétera, representan una buena base indispensable para comprender los avances de la electrónica y de las técnicas de los ordenadores y de la computación. *El núcleo atómico y las transformaciones nucleares* se tratan, de forma no exhaustiva, en los capítulos 13 y 14. (L. B.)

PROBLEMAS RESUELTOS DE ALGEBRA, por Emilio Espada. Editorial Eunibar; Barcelona, 1978-83. El título y el prólogo pueden originar alguna confusión, pues no se trata de una obra que recoja problemas correspondientes a los programas de álgebra del primer ciclo de la carrera de exactas (teoría de ecuaciones, álgebra conmutativa, etcétera), sino de una colección de problemas de álgebra lineal. Consta de doce capítulos repartidos en dos volúmenes. (El autor promete un tercer volumen que constará de cuatro capítulos.) Cada capítulo consta, a su vez, de una colección de problemas resueltos y de un breve resumen de las definiciones y los teoremas implicados. El nivel de los problemas propuestos, excepción hecha de los del último capítulo, es el que corresponde a los programas de COU y del primer curso de cualquier carrera de ciencias.

El capítulo 1 se dedica a la teoría de conjuntos. Se ocupa luego de las leyes de composición y las álgebras de Boole. Este capítulo es breve y elemental, y la única álgebra de Boole que se estudia con detalle es la de los circuitos eléctricos. El capítulo 3, sobre teoría de grupos, presenta problemas resueltos con gran detalle sobre la definición de grupo, subgrupo, etc., y sobre las consecuencias del teorema de Lagrange. En los últimos problemas de este capítulo se inicia el estudio de los grupos de permutaciones. En el cuarto nos ofrece problemas elementales sobre los anillos, ideales y cuerpos; al final del capítulo se muestran algunas de las aplicaciones de estos conceptos a las congruencias de números enteros. Los capítulos 5 y 6 se centran en los espacios vectoriales, aplicaciones lineales, y matrices. Además de los problemas que ejemplifican conceptos fundamentales del álgebra lineal (bases, independencia lineal, dimensión, etcétera), hallamos aquí una colección muy completa de problemas sobre aplicaciones lineales y sobre el estudio de sus imágenes y sus núcleos.

El séptimo nos trae la resolución de sistemas de ecuaciones lineales mediante el método de Gauss, e incluye sistemas de ecuaciones lineales que dependen de uno o de varios parámetros. En el siguiente aborda el espacio dual, cálculo de bases duales y análisis de las propiedades de la aplicación traspuesta de las aplicaciones lineales, aunque no se dan las aplicaciones del espacio dual a los problemas de incidencia.

Versa el capítulo 9 sobre tensores, cálculo de su expresión en las bases usuales, simetrización y antisimetrización de tensores y producto exterior de tensores antisimétricos. Pasa luego a los determinantes y a su aplicación a la resolución de sistemas de ecuaciones lineales. Un gran acierto del autor es introducir desde el principio el determinante de un endomorfismo y observar que, así como es posible definir el determinante de los endomorfismos, el determinante de *n* vectores (en un espacio de dimensión *n*) sólo tiene sentido cuando se ha prefijado una base. En cambio, cuando el autor introduce el producto escalar y el producto vectorial define ambos por su expresión en coordenadas (por supuesto cartesianas, lo cual dificulta notablemente su cálculo en los sistemas de referencia que no sean ortonormales), a pesar de que el resto de las definiciones del libro son intrínsecas y los conceptos necesarios para definir el producto escalar y el producto vectorial (tensores, orientaciones, etcétera) se han introducido con anterioridad.

El capítulo 11 está dedicado a los polinomios y a las fracciones racionales. Es conveniente advertir al lector de algunos errores: en el resumen teórico

correspondiente se introducen los conceptos de máximo común divisor y mínimo común múltiplo para polinomios con coeficientes en un anillo conmutativo arbitrario, aunque la existencia y las propiedades que se afirman sólo se verifican cuando el anillo de coeficientes es un cuerpo. Este error no es el olvido de una hipótesis, porque en los problemas se utilizan estos conceptos para polinomios con coeficientes enteros y con coeficientes en el anillo de los enteros de Gauss, de modo que es necesario saber que estos anillos son dominios de factorización única (concepto que no se introduce en el libro) y modificar la definición de máximo común divisor y mínimo común múltiplo. En este capítulo el estudio del anillo de los polinomios se aplica a las cuestiones de racionalización, de interpolación, y a la descomposición de las fracciones racionales en fracciones simples. También se realiza un estudio detallado de la estructura del anillo cociente del anillo de los polinomios por un ideal.

Muy completo y equilibrado es el capítulo 12: clasificación de endomorfismos. Se enseña el modo de clasificar los endomorfismos, de hallar la descomposición del espacio en monógenos irreducibles y de calcular el polinomio anulador, la forma canónica y las bases de Jordan de los endomorfismos. En un apéndice se muestra, además, cómo la clasificación de endomorfismos permite resolver las ecuaciones diferenciales ordinarias lineales homogéneas con coeficientes complejos, y los sistemas lineales homogéneos de ecuaciones diferenciales ordinarias con coeficientes complejos. Este capítulo contrasta vivamente con el carácter elemental de los restantes; su desarrollo es acertado. Uno de los mayores defectos de los restantes capítulos es no haber utilizado los problemas para exponer los ejemplos donde la teoría en cuestión se aplica corrientemente y desarrollar a su vez las aplicaciones más importantes de los conceptos que se introducen.

El autor pone el acento sobre los cálculos en coordenadas, en detrimento de los cálculos intrínsecos. Tan importante es saber calcular en los diferentes sistemas de coordenadas como evitar los cómputos en coordenadas cuando es posible sustituirlos por razonamientos intrínsecos más sencillos y claros. (J. A. N.)

Ciencia y Cultura al Servicio del Hombre. Dirigido por Alejandro Llano. Editorial Dossat; Madrid, 1982. Este volumen recoge la transcripción de cuatro mesas redondas que tuvieron lugar en el Colegio Mayor Zurbarán de Madrid en el otoño de 1981. Participaron 14 españoles y dos franceses, actuando como moderador Alejandro Llano, decano de la Facultad de Filosofía de la Universidad de Navarra.

"La sociedad se preocupa mucho ahora por atribuir un papel moral a la ciencia. Es un esfuerzo inútil ya que el hallazgo científico, por sí mismo, no es bueno ni malo: lo que ocurre es que ese hallazgo no está en el vacío, está hecho por hombres y para los hombres. La ciencia brinda a los hombres una cosa hermosa: el conocimiento, y una cosa difícil: la responsabilidad". Estas palabras de Julio R. Villanueva, en el prólogo, enmarcan en buena medida estos diálogos, en los que se alude a la tensión frecuente entre un progreso científico intrínsecamente beneficioso y unas consecuencias negativas que se estudian en sus raíces: el incremento de la desigualdad social, los desastres ecológicos, la utilización bélica de los conocimientos científicos, los riesgos de la energía nuclear y de la manipulación genética, las ideologías pseudocientíficas.

"La ciencia actual entre el temor y la esperanza" es el título del primer coloquio, en el que intervienen los catedráticos Alonso Santos (energía nuclear), Carpena Artés (química agrícola), Julio R. Villanueva (microbiología) y el profesor Díez Hochleitner. El debate se centra en las expectativas y temores suscitados por el progreso científico de las últimas décadas, aportando numerosos datos y reflexiones de gran interés. Sólo cinco países consumen casi el 85 por ciento de los fondos mundiales que se dedican a la investigación; más de dos millones y medio de científicos y técnicos trabajan en los países desarrollados, mientras que unos trescientos mil lo hacen en los países en desarrollo. En la actualidad hay capacidad potencial para alimentar a cuarenta mil millones de personas, pero el problema del hambre no se resolverá mientras los países desarrollados no inviertan mucho más para hacer posible la investigación agrícola aplicada a los países que necesitan ser alimentados. En el curso del diálogo, Rodríguez Villanueva matiza los riesgos sociales de la reciente ingeniería genética, y sostiene una actitud de esperanza fundada en una consideración ética de la investigación.

En el segundo coloquio, con el título "La ciencia y el enigma del hombre", intervienen los catedráticos José María Macarulla (bioquímica), Sánchez del Río (física atómica y nuclear), Claude Bruaire (filosofía, de la Sorbona) y Pedro Rodríguez (teología dogmática). La pregunta: ¿qué nos puede decir la ciencia acerca del misterio de la humana condición? constituye el *leit-motiv* de la sesión. La cuestión así formulada pone de manifiesto la naturaleza propia del trabajo científico. "La ciencia –dice Sánchez del Río– es, básicamente, una descripción, y no una explicación. A pesar de que frecuentemente se dice que se explican las cosas, se trata en realidad de una descripción. La ciencia consiste en la formulación de unas cosas obvias de manera tal que las consecuencias que se deduzcan ya no sean obvias... En el fondo no responde a la pregunta del porqué, responde únicamente a la pregunta del cómo" (p. 52). En este sentido, Sánchez del Río y Macarulla dan cuenta de las descripciones científicas de la génesis del cosmos y de la génesis evolutiva de la vida, respectivamente; pero son el filósofo Claude Bruaire y el teólogo Pedro Rodríguez quienes abordan ese *plus* de realidad por el que el hombre trasciende su descripción científica.

Esta temática prosigue en el tercer coloquio, centrado en el tema "Ciencia y cultura ante la dignidad humana". Alejandro Llano plantea desde el comienzo el problema: "nunca como hasta ahora la dignidad humana se había proclamado con más fuerza, y nunca como hasta ahora se ha conculcado de manera más eficaz" (p. 63). Intervienen en el debate los catedráticos Larrodera López (planeamiento urbanístico), de la Cuesta Rute (derecho de la publicidad) y Cervós Navarro (neuropatología), quienes consideran desde distintos puntos de vista la manipulación de que el hombre puede ser objeto mediante la aplicación de los conocimientos científicos y técnicos. Emilio Larrodera denuncia la pérdida de la relación humana en la ciudad actual. José M.ª de la Cuesta llama la atención sobre el sentido de *mediación* de los medios de comunicación social. Jorge Cervós, de la Universidad Libre de Berlín, postula trascender los reduccionismos simplificadores, en línea con Eccles, en *El yo y su cerebro* (Editorial Labor, 1982): "la imponente cerebralización del hombre, que es un sustrato necesario para sus funciones psíquicas, no puede ser la causa de esas funciones psíquicas" (p. 82).

El último coloquio recogido en el libro, "Ciencia y fe", es especialmente ágil y polémico, como han puesto de re-

lieve los recientes cruces epistolares de las distintas posiciones en los medios de comunicación científica. Intervienen los profesores Gutiérrez Ríos (química inorgánica), Alvaro del Amo (genética), Mariano Artigas (filosofía de la ciencia) y el filósofo y escritor francés Christian Chabanis. Se registra el profundo cambio que de modo paulatino se ha producido en la mentalidad científica moderna: frente a posturas reduccionistas todavía influyentes en los comienzos de nuestro siglo, se ha tomado conciencia acerca de la especificidad de los métodos científicos, con sus posibilidades y limitaciones. Chabanis aporta una curiosa encuesta sobre el problema religioso en distintos profesionales de la ciencia. Mariano Artigas analiza el tránsito del antiguo cientificismo optimista al moderno, de corte pesimista, según el cual todo conocimiento sería conjetural, y muestra las insuficiencias de ambos planteamientos. Gutiérrez Ríos da una visión clara de las motivaciones y actitudes del investigador. Alvaro del Amo expone la situación actual de las teorías evolucionistas: "el antagonismo entre creación y evolución carece de sentido". (J. N.)

AN ALBUM OF FLUID MOTION, por Milton Van Dyke. Parabolic Press; Stanford, 1982. Visualizar lo que ocurre en un fluido (líquido o gas) en movimiento ha sido y es un arte, una ciencia y una técnica en continua evolución desde que el hombre inició su estudio experimental de la naturaleza y quiso verificar que lo que imaginaba sucedía era muy diferente de la realidad. Actualmente se pueden ver los movimientos convectivos, las corrientes laminares y las turbulentas, es decir, la distribución de velocidades en un fluido, así como la distribución de temperaturas o de partículas de diversos tamaños, distribuciones de presión y de gradientes de esas magnitudes. En cada punto, se puede ver, por ejemplo, el valor de la temperatura del fluido en movimiento, así como su tendencia instantánea hacia otro valor superior o inferior.

Tales visualizaciones empezaron siendo de interés cualitativo, fenomenológico, descriptivo. Hoy, la precisión alcanzada y la posibilidad de obtener y procesar datos y reconocer estructuras digitalizando imágenes estáticas y dinámicas con ordenadores y otros equipos ópticos y electrónicos de rapidísima respuesta permiten el estudio cuantitativo de las propiedades de las corrientes fluidas. La representación óptica es valiosa para el investigador y el docente. La visualización de corrientes fluidas ha sido crucial en aquellas ocasiones en que no hemos sido capaces de construir, de antemano, la solución teórica del problema. Contemplando la solución, ha sido posible recrearla teóricamente y así se ha avanzado en el problema. Y si no, que se lo pregunten a los ingenieros aeronáuticos, navales, de automóviles, de puentes, etcétera. Cuán poco entendemos aún cómo se genera y desarrolla la turbulencia, o cómo funciona la aerodinámica de objetos en régimen turbulento; sin embargo, hemos sido capaces de construir aeronaves maravillosas que llegan a regiones hasta hace muy poco ignotas.

El presente "Album of fluid motion" es el primer libro exclusivamente dedicado a presentar fotos de corrientes convectivas. Es la labor de Milton Van Dyke, de la Universidad de Stanford, célebre por sus trabajos teóricos en mecánica de fluidos (particularmente, en el uso de la técnica de perturbaciones singulares). Es una obra familiar por tres razones: el libro es producto de la colaboración con su esposa, lo ha publicado él mismo (Parabolic Press es editorial suya) y, por último, la colección de fotos proviene de las aportaciones voluntarias y gratuitas de la gran *familia* de cuantos cultivamos ese campo, colegas y amigos que Milton tiene por doquier. Las dos primeras razones han conducido a una producción esmerada con ese cuidado que sólo suele darse a lo que se ama. La tercera condujo a una impresionante colección científica, histórica y artísticamente hablando.

Artísticamente, nos encontramos con un libro maravilloso, donde figuras enrevesadas aparecen junto a otras simples en una amalgama que da idea de la complejidad de la materia tratada. Lástima que sólo hay imágenes en blanco y negro. Hay aquí fotos de objetos más bellos que muchas de las pinturas que se exhiben en museos. Pero en el Album de Milton la obra, aunque generalmente forzada, creada, en el laboratorio es de la propia naturaleza: el hombre sólo capturó algunos instantes de belleza.

Científicamente, el Album nos hace recorrer las más relevantes posibilidades de movimientos convectivos: corrientes con muy baja velocidad (para pequeños valores, inferiores y muy inferiores a la unidad, del número de Reynolds, que es la razón de la velocidad y de la dimensión del objeto a la viscosidad del fluido), con velocidades crecientes hasta alcanzar valores muy elevados (con números de Reynolds desde la decena hasta varios centenares de millar y para objetos de diversa forma). Maravillosas ilustraciones muestran corrientes laminares, sencillas, y turbulentas, complejas, donde lo caótico y lo estructurado coexisten.

La separación de corrientes fluidas alrededor de objetos sólidos, torbellinos en el fluido, olas (ondas), estructuras periódicas tanto en el interior como en la superficie de un fluido en movimiento, ondas de choque, corrientes supersónicas y subsónicas (cuando la velocidad del fluido es próxima, superior o inferior, a la velocidad del sonido en dicho fluido), etcétera, son fenómenos cuya ilustración se nos ofrece aquí.

Todas las fotos llevan un pie de figura suficiente como para que el lector consiga hacerse una idea de lo que hasta ahora hemos entendido del fenómeno captado por la imagen. Los fenómenos se agrupan en once capítulos, a los que se añaden dos páginas de índice analítico y dos páginas de referencias que lo convierten en un instrumento utilísimo para el estudioso de los fluidos.

Verdadero archivo histórico, se recogen fotos de excepcional valor: dos del movimiento supersónico de una bala tomada por Ernst Mach en 1888 y 1893, respectivamente, en su laboratorio de Praga (la primera aumentada treinta veces para llenar una página a partir de un negativo milimétrico).

No pretende ser libro de texto alguno. Y, por otro lado, cabe considerarlo una joya de arte. Picasso, Goya, Rubens, Zurbarán, El Greco,... son pintores cuyas obras, se reconoce, deben ser familiares a cualquier ciudadano alfabetizado, si no culto. En las postrimerías del siglo XX parece razonable (si no exigible) tener conocimiento de fenómenos científicos y tecnológicos del entorno en consonancia con el impacto que tanto la ciencia como la técnica tienen sobre nuestra vida.

Por último, de entre las casi trescientas fotos, España está representada por una ilustración del trabajo de investigadores de la Universidad Nacional de Educación a Distancia (foto n.º 141, pág. 83) relativa a la convección natural inducida por el acoplamiento entre fuerzas de flotación de Arquímedes y efectos de tensión superficial, aparecida como primicia en INVESTIGACIÓN Y CIENCIA (octubre de 1980). Hay también una bellísima foto de turbulencia (foto n.º 177, pág. 102) extraída de la tesis doctoral de otro español. (M. G. Velarde.)

Bibliografía

Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:

RESPUESTA PRECOZ DE CONTRAGOLPE

LAUNCH UNDER ATTACK TO REDRESS MINUTEMAN VULNERABILITY? Richard Garwin en *International Security*, vol. 4, n.º 3, págs. 117-139; invierno, 1979-80.

NUCLEAR DECAPITATION. John D. Steinbruner en *Foreign Policy*, n.º 45, págs. 16-28; invierno, 1981-82.

A COMPUTER SIMULATION OF LAUNCH PROCEDURES FOR A SQUADRON OF MINUTEMAN II ICBMs. Paul Morawski en *Simulation*, vol. 41, n.º 1, páginas 16-24; julio, 1983.

EFECTOS ATMOSFERICOS DE EL CHICHON

STRATOSPHERIC SULFURIC ACID FRACTION AND MASS ESTIMATE FOR THE 1982 VOLCANIC ERUPTION OF EL CHICHON. D. J. Hofmann y J. M. Rosen en *Geophysical Research Letters*, vol. 10. n.º 4, págs. 313-316; abril, 1983.

STRATOSPHERIC AEROSOL MASS AND LATITUDINAL DISTRIBUTION OF THE EL CHICHON ERUPTION CLOUD FOR OCTOBER 1982. P. M. McCormick y T. J. Swissler en *Geophysical Research Letters*, vol. 10, n.º 9, págs. 877-880; septiembre, 1983.

CLIMATIC EFFECTS OF THE ERUPTION OF EL CHICHON. Número especial de *Geophysical Research Letters*, vol. 10, n.º 11, págs. 989-1060; noviembre, 1983.

COLISIONES A ALTA ENERGIA ENTRE NUCLEOS ATOMICOS

NUCLEUS-NUCLEUS COLLISIONS: PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON NUCLEUS-NUCLEUS COLLISIONS, MICHIGAN STATE UNIVERSITY, SEPTEMBER 26-OCTOBER 1, 1982. Dirigido por G. F. Bertsch, C. K. Gelbke y D. K. Scott. North-Holland Publishing Co., 1983.

ANOMALONS AS PINUETS BOUND TO NUCLEAR FRAGMENTS: A POSSIBLE EXPLANATION. Wm. C. McHarris y J. O. Rasmussen en *Physics Letters*, vol. 120B, n.º 1, 2, 3, págs. 49-53; 6 de enero de 1983.

ARROZ

TROPICAL CROPS: MONOCOTYLEDONS 1. J. W. Purseglove. John Wiley & Sons, Inc., 1972.

AFRICAN CEREALS: *ELEUSINE*, FONIO, BLACK FONIO, TEFF, *BRACHIARIA*, *PASPALLUM*, *PENNISETUM* and AFRICAN RICE. R. Porteres en *Origins of African Plant Domestication*, dirigido por Jack R. Harlan, Jan M. J. de Wet y Ann B. L. Stemler, Mouton Publishers, 1976.

THE ORIGIN, EVOLUTION, CULTIVATION, DISSEMINATION, AND DIVERSIFICATION OF ASIAN AND AFRICAN RICES. T. T. Chang en *Euphytica*, vol. 25, págs. 425-441; 1976.

PRINCIPLES AND PRACTICES OF RICE PRODUCTION. Surajit K. De Datta. John Wiley & Sons, Inc., 1981.

LAGARTIJAS UNISEXUALES

EVOLUTION OF PARTHENOGENETIC SPECIES OF REPTILES. Charles J. Cole en *Symposium on Intersexuality in the Animal Kingdom*, dirigido por Rudolf Reinboth. Springer-Verlag, 1975.

CHROMOSOME INHERITANCE IN PARTHENOGENETIC LIZARDS AND EVOLUTION OF ALLOPOLYPLOIDY IN REPTILES. Charles J. Cole en *Journal of Heredity*, vol. 70, n.º 2, págs. 95-102; 1979.

PARTHENOGENETIC REPRODUCTION IN LIZARDS: HISTOLOGICAL EVIDENCE. Laurence H. Hardy y Charles J. Cole en *Journal of Morphology*, vol. 170, n.º 2, págs. 215-237; 1981.

CONTROL DE LA SINTESIS DE RIBOSOMAS

REGULATION OF RIBOSOMAL PROTEIN SYNTHESIS EN *ESCHERICHIA COLI* BY SELECTIVE mRNA INACTIVATION. Ann M. Fallon, C. Sue Jinks, Geneva D. Strycharz y Masayasu Nomura en *Proceedings of the National Academy of Sciences of the United States of America*, vol. 76, n.º 7, págs. 3411-3415; julio, 1979.

IN VITRO EXPRESSION OF *ESCHERICHIA COLI* RIBOSOMAL PROTEIN GENES: AUTOGENOUS INHIBITION OF TRANSLATION. John L. Yates, Ann E. Arfsten y Masayasu Nomura en *Proceedings of the National Academy of Sciences of the United States of America*, vol. 77, n.º 4, págs. 1837-1841; abril, 1980.

REGULATION OF RIBOSOME BIOSYNTHESIS IN *ESCHERICHIA COLI*. Masayasu Nomura, Sue Jinks-Robertson y Akiko Miura en *Interaction of Translational and Transcriptional Controls in the Regulation of Gene Expression*, dirigido por M. Grunberg-Manago y B. Safer, Elsevier Science Publishing Company, Inc., 1983.

EMPAQUETAMIENTO DE ESFERAS

TWELVE GEOMETRIC ESSAYS. H. S. M. Coxeter. Southern Illinois University Press, 1968.

LAGERUNGEN IN DER EBENE, AUF DER KUGEL UND IM RAUM. L. Fejes Tóth. Springer-Verlag, 1972.

THE LEECH LATTICE, SPHERE PACKINGS, AND RELATED TOPICS. J. H. Conway y N. J. A. Sloane, Springer-Verlag, 1984.

MINEROS Y METALURGICOS ANTIGUOS EN HUELVA

MINEROS ANTIGUOS ESPAÑOLES. A. Blanco y J. M. Luzón en *Archivo Español de Arqueología* XXXIX, páginas 73-83; 1966.

PRE-ROMAN SILVER MINERS AT RIOTINTO. A. Blanco y J. M. Luzón en *Antiquity* XLIII, págs. 124-131; 1969.

ANCIENT SLAGS IN THE SOUTH WEST OF THE IBERIAN PENINSULA. L. U. Salkield en *La minería hispana e iberoamericana*, I, págs. 85-98; León, 1970.

ANCIENT COPPER MINING AND SMELTING AT CHINFLON (HUELVA, SW SPAIN). B. Rothenberg y A. Blanco en *Scientific Studies in Early Mining and Extractive Metallurgy* (British Museum, Occasional Papers). London, 1980.

EXPLORACIÓN ARQUEOMETALÚRGICA DE HUELVA. A. Blanco Freijeiro y B. Rothenberg. Edit. Labor; Barcelona, 1981.

JUEGOS DE ORDENADOR

THE $3x + 1$ PROBLEM AND ITS GENERALIZATIONS. J. C. Lagarias. De próxima aparición en *The American Mathematical Monthly*.

TALLER Y LABORATORIO

LIGHT AND COLOR IN NATURE AND ART. Samuel J. Williamson y Herman Z. Cummins. John Wiley & Sons, Inc., 1983.